

Consortium for International Earth Science Information Network



1N-42-CR
5259
p. 65

A Report to:

THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
GODDARD SPACE FLIGHT CENTER

Code 285
Greenbelt, MD 20771 USA

LETTER CONTRACT NAS5-32632
Contract Data Requirements List Item 12

Submitted by:

Consortium for International Earth Science Information Network

2250 Pierce Road
University Center, Michigan 48710 USA
Phone: (517) 797-2700
Fax: (517) 797-2622

Principal Investigator:
Dr. Robert S. Chen
Program Manager



October 1995

N96-12236

Unclass

G3/42 0071420

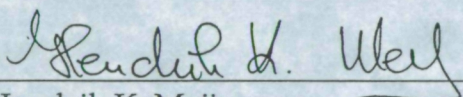
(NASA-CR-199581) SEDAC INFORMATION
GATEWAY PLAN V(1) Annual Report, 1
Oct. 1994 - 30 Sep. 1995
(Consortium for Intern. Earth Sci.
Info. Network) 65 p

Consortium for International Earth Science Information Network
Saginaw, Michigan

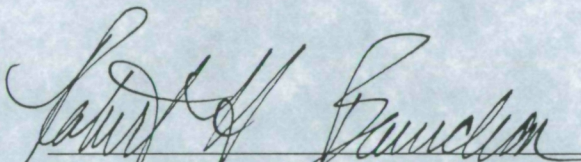
SEDAC Information Gateway Plan V(1)

Prepared Under Letter Contract NAS5-32632

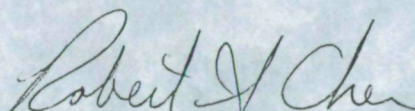
APPROVED BY:

 9-29-95

Dr. Hendrik K. Meij
SEDAC Scientist Date

 9-29-95

Dr. Robert Bourdeau
SEDAC System Engineer (Acting) Date

 9/29/95

Dr. Robert S. Chen
SEDAC Manager Date

October 1995

Preface

This annual update to the *Information Gateway Plan* is submitted as required by Section F (Deliveries or Performance) of Contract NAS5-32632 for the Development and Operation of the Socioeconomic Data and Applications Center (SEDAC) in fulfillment of CDRL-13. This annual update of the Plan incorporates changes recommended by the SEDAC User Working Group (UWG) and reflects comments and suggestions from users, collaborators, and the Contracting Officer Technical Representative (COTR).

The Information Gateway Plan is a concise and specific plan that outlines SEDAC activities and services in support of the Earth and social sciences and other user communities. The SEDAC Information Gateway provides interdisciplinary access to socioeconomic and physical science data and information resources held by SEDAC and numerous other institutions and networks around the world.

This version of the Plan was prepared by the SEDAC Information Gateway Team with the guidance of the SEDAC Project Scientist and the SEDAC Manager. It has been reviewed and approved by the SEDAC UWG.

Questions or proposed changes
should be addressed to:

SEDAC Manager
CIESIN
2250 Pierce Road
Saginaw, Michigan 48710 USA
Tel. 517-797-2611

Executive Summary

This Information Gateway Plan describes how the Socioeconomic Data and Applications Center (SEDAC) at the Consortium for International Earth Science Information Network (CIESIN) is responding to SEDAC's mission to serve as a two-way "Information Gateway" between the socioeconomic and Earth science data and information domains. The SEDAC Information Gateway effort is a primary means by which the Earth Observing System Data and Information System (EOSDIS) can link meaningfully with a broad range of social science data sources and users in ways that lead to tangible benefits to the American people.

The Plan describes the areas of research of earth scientists and socioeconomic scientists where interchange of data and information is most needed. It sets guidelines for the continued development of SEDAC's directory of social science datasets and establishes priorities for efforts to make data held by SEDAC or accessible through SEDAC available to the user community. The Plan also describes the means by which the SEDAC user community can access information products specified by the SEDAC Data and Applications Development Plan (DADP). Among other major activities, SEDAC will continue to enhance and operate a directory capability, interoperable with the Global Change Master Directory, that provides the socioeconomic community with information about Earth science products and the Earth science research community with information about socioeconomic data. The Information Gateway also serves as a unique and powerful access pathway for a wide range of users and potential users of socioeconomic and earth science data, including especially remote sensing data.

SEDAC recognizes the vital role that the emerging National Information Infrastructure (NII) has on improving access to scientific data and information and the application of such data and information in day-to-day decision making. This updated plan places a renewed emphasis on the use of the Internet, which is today's operational component of the NII. SEDAC is employing available capabilities, such as the World Wide Web (WWW) and file transfer protocol (ftp), and is developing special tools as necessary to deliver information products to users throughout the Internet. These tools include the Gateway, an innovative award-winning system that allows SEDAC users to identify data of interest.

During its second year, SEDAC will continue to pursue a three-fold technical strategy to accomplish its Information Gateway mission:

- *Development of a catalog of socioeconomic and environmental data and information resources focused on human interactions in the environment.* This catalog will extend the EOSDIS and Global Change Data and Information System (GCDIS) efforts to document natural science data into the socioeconomic data domain. The cataloging effort will conform to recommended EOSDIS guidelines. The catalog will be made available through several different access methods and channels to maximize accessibility and usability throughout the targeted user communities.

- *Building cooperative relationships with data and information providers and archives from around the world.* This effort takes advantage of CIESIN's Information Cooperative, initiated under previous NASA Grant efforts at CIESIN. The Information Cooperative recognizes the distributed nature of data and information resources, the rapidly improving capabilities for making data available reliably and cheaply over the Internet, the importance of keeping control of data with data sources or with strong data institutions, and the benefits of institutional collaboration in developing value-added, interdisciplinary data products and services.
- *Provision of online services and user assistance to complement the above.* These services significantly improve the ability of users to access and utilize data products and applications provided by SEDAC and other linked organizations. Services include direct access to data and information resources at local and distributed archives, an order and retrieval system, information guides, information resource discovery tools, data analysis and processing tools, and user services support. Coordination and integration of these services occurs primarily through a SEDAC World Wide Web site.

This second version of the Plan has been prepared under the guidance of the SEDAC Project Scientist to ensure compatibility with the SEDAC DADP and with the Science Data Plan currently being updated. This update reflects guidance from the SEDAC User Working Group, feedback from users, and requirements arising from SEDAC's data and applications development activities.

Contents

Preface	i
Executive Summary	iii
1.0 Introduction	1
1.1 Past Year Accomplishments	2
1.2 Objectives of the Mission to Planet Earth Program.....	4
1.3 SEDAC's Role in Mission To Planet Earth	5
2.0 The Nature of Global Change Research: Human Dimensions Context ...	8
3.0 Information Gateway Mission and Requirements	11
3.1 Definition.....	12
3.2 User Community	13
3.3 Requirements for Access to Socioeconomic Data and Information Resources	13
3.4 Requirements for Access to Physical Science Resources	32
3.5 Requirements for Integration of Physical and Socioeconomic Data.....	35
4.0 Information Gateway Implementation Mechanisms	38
4.1 Information Model	38
4.2 National Information Infrastructure	42
4.3 Information Cooperative.....	42
4.4 The SEDAC Archive	46
5.0 Information Gateway System Components High Level Requirements...	49
5.1 Data and Information Access System	50
5.2 Data Archive and Distribution System	51
5.3 Data Processing and Analysis System	53
5.4 User Services System	53
References	55
Acronyms and Abbreviations	57

Figures and Tables

Figures

1. The Social Process Diagram 15
2. Simplified version of the "Bretherton" diagram 33

Tables

1. Relation between SEDAC Science Data Domains and U.S. NRC reports concerning the human dimensions of global environmental change 17
2. Relation between SEDAC Science Data Domains and USGCRP and HDP reports concerning the human dimensions of global environmental change 18
3. Example Information Cooperative Partners with significant holdings corresponding to SEDAC Science Data Domains 47
4. Information Cooperative National Partners 48

1.0 Introduction

The Socioeconomic Data and Applications Center (SEDAC) is one of nine Distributed Active Archive Centers (DAACs) in the Earth Observing System Data and Information System (EOSDIS) that engage in processing, cataloging, archiving, and distributing data derived from space-based platforms, including Earth Observing System and Earth Probe flight missions undertaken as part of NASA's Mission to Planet Earth (MTPE) program. One SEDAC priority mission is to support the MTPE program by contributing to its goal of translating scientific understanding into tangible benefits to the American people by developing new policy-oriented information products that synthesize Earth science and socioeconomic data and by providing the resulting operational data and information services to the public. A second SEDAC mission of equal priority is to serve as a two-way "Information Gateway" between the socioeconomic and Earth science data and information domains.

This Information Gateway Plan (IGP) addresses SEDAC's second mission. The Information Gateway mission is closely coupled with SEDAC's first mission but also reflects the critical need for interdisciplinary collaboration and exchange across the natural and social science communities. The IGP has therefore been developed in conjunction with the SEDAC Data and Applications Development Plan (DADP), which describes the data and applications development effort to be carried out by SEDAC. The IGP attempts to address the broader needs for interaction between the natural and social science communities, based on past user feedback and experience, guidance from the SEDAC User Working Group (UWG), and other relevant inputs. It also deals with the related needs of the broader set of actual and potential users of data and applications provided by SEDAC and other components of EOSDIS.

The IGP outlines the general nature of scientific research in the human dimensions of global environmental change and highlights specific recommendations by the scientific community, both nationally and internationally, in the last few years. Based on these recommendations, the Plan defines a set of eight interdisciplinary data "domains" in which SEDAC has initially concentrated its data cataloging efforts. These domains characterize the major classes of data useful in data integration efforts by SEDAC and the scientific and data management communities at large.

The IGP and the DADP together provide key inputs into the SEDAC Science Data Plan (SDP). The SDP details the specific datasets that SEDAC will provide, the institutional linkages to be developed or strengthened, and other information pertinent to how data products are archived, accessed, and managed.

The IGP, along with the other key SEDAC plans, is intended as a "living" document. Both scientific knowledge and policy concerns about global environmental change are evolving rapidly. Many data sources and data and information networks are changing

in response to rapid changes in technology and economic conditions. Needs, priorities, and opportunities will change over time as well. It is therefore critical that SEDAC remain flexible and keep a long-term perspective on data and information needs. By updating and revising the IGP and other plans based on careful interactions with the scientific and data management communities, key user groups, and NASA management and the other EOSDIS DAACs, SEDAC can maintain and improve the relevance and utility of its products and services while at the same time ensuring coordination and sound management of its resources.

This year's update to the IGP draws heavily on lessons learned during the past year. Recommendations by the SEDAC UWG have been used to delineate areas of focus within science domains for data identification purposes. Experience with actual users has generated ideas for enhancements to SEDAC Information Gateway services. Interactions with CIESIN's Information Cooperative partners have helped to define strategies for strengthening SEDAC's distributed catalog system and other cross-institution products and services. Finally, the SEDAC Information Model has been modified to reflect ongoing refinements of metadata standards.

1.1 Past Year Accomplishments

During its first year of activity, SEDAC has developed a range of information products and services as a result of Information Gateway efforts. These include:

- 1) *A Gateway system* which provides distributed search-and-order capabilities across widely scattered directory and inventory servers using a common user interface. Beta versions of the Gateway clients are now available via the Internet for UNIX X-Windows workstations and for Microsoft Windows-compatible personal computers. A character-based interface for directory-level searching is also available. Servers are functioning that provide access to a variety of directory- and inventory-level metadata including the GCMD. The Gateway system has recently won first prize in the Environment, Energy & Agriculture category of the prestigious 1995 Computerworld Smithsonian Awards Program and will be part of the Smithsonian's permanent research collection and a public exhibit at the National Museum of American History.
- 2) *A collection of human dimensions metadata.* SEDAC's metadata collection now contains more than 480 directory and 1,700 inventory entries. A total of 43 directory entries have been submitted to the GCMD. Progress in metadata development is more fully documented in the Metadata Development Report (June 1995), a subtask 5.2 milestone.
- 3) *Memoranda of Understanding (MoUs) or other agreements* with more than a dozen different organizations (partners in CIESIN's Information Cooperative) to share metadata, provide access to data, and/or collaborate in other ways. Key partners relevant to SEDAC applications development include the International Programs Center of the U.S. Bureau of the Census (formerly the Center for International Research), EROS Data Center (EDC), the World Conservation Union (IUCN), the

National Institute on Public Health and Environmental Protection (RIVM) in the Netherlands, the United Nations Environment Programme/Global Resource Information Database (UNEP/GRID), the World Conservation Monitoring Centre (WCMC), the World Resources Institute, and key institutions in China.

- 4) *Information Guides available over the Internet.* SEDAC supports several guide types, including dataset guides, organizational guides, and thematic guides. A total of 43 information guides are currently available.
- 5) *A Human Dimensions Kiosk* for dissemination of pre-published papers and articles on human dimensions of global environmental change.
- 6) *SEDAC World Wide Web pages.* The World Wide Web (WWW) on the Internet provides a consistent point of entry into SEDAC's interdisciplinary data and information resources and other tools and services.
- 7) *Training Programs.* SEDAC provided assistance in the training of Information Cooperative node administrators from China and eastern Europe in metadata management, in order to promote interoperability of catalog services.

Recently, CIESIN has been nominated by the U.S. National Research Council to become a World Data Center-A for Human Interactions in the Environment, based in part on SEDAC Information Gateway services.

In support of these Information Gateway efforts, SEDAC has also designed and made significant progress in implementing the necessary infrastructure to provide efficient and reliable services to users. This infrastructure includes:

- the basic hardware and software systems for managing SEDAC data and information resources and services;
- facilities for housing SEDAC equipment, personnel, and data;
- working relationships and linkages with key user communities and groups, including the SEDAC UWG;
- basic design concepts, implementation procedures, and other plans as thoroughly documented in various contract deliverables; and
- highly qualified personnel and other individual and institutional contributors that provide the range of capabilities necessary to develop and implement innovative services and products, properly manage data and information resources, and support user needs.

For further information about recent SEDAC progress and user activity, please see the *SEDAC Annual Progress Report*, the *User Model Report*, and other SEDAC documents.

1.2 Objectives of the Mission to Planet Earth Program

During the past year, NASA and the MTPE program have made significant progress in refining and articulating their goals and strategies. In particular, the *Mission to Planet Earth Strategic Enterprise Plan 1995-2000* (May 1995) states:

The overall mission of NASA's Mission to Planet Earth Enterprise is to develop understanding of the total Earth system and the effects of natural and human-induced changes on the global environment (NASA, 1995a: 5).

The Plan goes on to identify six goals:

- 1) Increase scientific understanding of the Earth as an integrated environmental system and of its vulnerability to natural variations and human influences;
- 2) Observe and characterize the entire Earth system and make resultant data widely available;
- 3) Contribute to wise and timely national and international environmental policy;
- 4) Foster the development of an informed and environmentally aware public;
- 5) Expand and improve the scientific return on investment by seeking and using advanced, cost-effective engineering and scientific techniques having potential for multiple uses;
- 6) Ensure the availability of human and physical resources necessary to achieve the MTPE mission over the long term.

A recent National Research Council (NRC) review of the U.S. Global Change Research Program (USGCRP) and NASA's Mission to Planet Earth/Earth Observing System recommends that:

The USGCRP as a whole, and NASA's Mission to Planet Earth (MTPE) Program specifically, should maintain a **science-driven** approach to observation and information technology that employs current technology while investing in the development of new technology with clear applications to support the program's specific scientific priorities. (NRC, 1995: 20-21).

The report goes on to point out that MTPE missions "will contribute to practical applications such as natural hazards mitigation, water resources management, and food and fiber production, as well as advances in the Earth sciences" (NRC, 1995: 23). In particular, the NRC recommends:

The capabilities of MTPE/EOS should be exploited fully via enhanced public access to the information products. (NRC, 1995: 23).

MTPE science priorities for 1995-2000 are expected to focus in four areas:

- Atmospheric ozone;
- Seasonal-to-interannual climate prediction;
- Long-term climate variability; and
- Land cover change and global productivity.

The goals of the MTPE stratospheric ozone research effort include documentation of trends and patterns of UV radiation at the Earth's surface, provision of a monitoring strategy for stratospheric ozone for the 21st century, and formal assessment of the state of scientific understanding of potential impacts of human actions and technologies on stratospheric ozone.

In the area of long-term climate variability, the MTPE program seeks among other things to understand causes of climate change by determining natural and human-related forcing mechanisms and the feedbacks that govern climate variability, to forecast predictable aspects of multi-year climate on a region-by-region basis, to assess formally the state of scientific understanding of climate change on a regular basis, and to develop a strategy for long-term monitoring of selected climate forcing and diagnostic variables.

The MTPE land-cover/land-use change program addresses four areas: 1) land-cover conversion; 2) land-use intensification; 3) verification; and 4) modeling. Program goals include the identification of the current distribution of land-cover types and their conversion to other types; improved understanding of the consequences of intensified management of agricultural and agroforestry systems in the tropics and sub-tropics; measurement of the long-term *in situ* degradation of forest ecosystems; improved understanding of the consequences of international agreements; and investigation of forestry and agriculture in temperate and boreal climates.

As part of its efforts to restore its research and analysis program, the MTPE program will support basic research to evaluate the potential of Lewis and Clark satellite data in applications to natural hazard risk assessment and mitigation, high precision forestry and agriculture, urban planning, and biodiversity studies. It will also contribute to an interagency effort to improve understanding and remote sensing methods related to water resource and natural hazard issues in the mid-continental U.S. Finally, it will support studies of critical land processes that affect the vulnerability of ecosystems and agriculture to weather and climate variations, focusing in particular on Mexico, China, and India.¹

1.3 SEDAC's Role in Mission To Planet Earth

SEDAC has a clear role to play in the MTPE program. For example, under Goal 3 of the Strategic Enterprise Plan cited above, NASA identifies objective 3.1, to "ensure the availability of MTPE data and understanding for policy makers". It points out that "one

¹ D. Butler, Office of Mission to Planet Earth, personal communication, September 1995.

of the nine DAACs is dedicated to social and economic sciences data, " i.e., SEDAC. It also outlines a strategy for facilitating "the increased understanding of global change issues among policy makers through the provision of sample user application scenarios by 1997" (NASA, 1995a: 23).

Under objective 3.2, NASA indicates that MTPE will "support all formal integrated assessments on the global environment beginning in 1995." The Plan specifically notes that "MTPE scientists have a responsibility to participate in national and international global change assessments," including those undertaken by the World Meteorological Organization, United Nations Environment Programme, and the Intergovernmental Panel on Climate Change (NASA, 1995a: 24).

Under objective 3.3, NASA recognizes the need for "integrated models" that:

...capture a sufficient level of detail available in individual arena models, and couple them in a manner that does not compound the errors from individual ones, or create new errors. The user must be able to run specified alternate assumption sets and meaningfully compare the resultant states predicted by the model (NASA, 1995a: 24).

As a strategy for meeting this objective, NASA indicates it will undertake to "apply at least one integrated model of the global atmosphere, oceans, and land surface incorporating space-derived data by 1998."

Clear roles for SEDAC are also evident under other goals. For example, as part of Goal 2's aim to make data widely available, SEDAC and the other DAACs not only carry out "responsibilities for processing, archiving, and distributing EOS data and related data and for providing a full range of user support", but also "support global change researchers whose needs cross traditional discipline boundaries" (NASA, 1995a: 22). These efforts to make data more available tie in closely with Goal 4, which includes the objectives of making information and assessments available to the general public and the educational community. The latter goal includes efforts to support curriculum development, put a MTPE "home page" in place on the Internet, and reach out to teachers and students. In addition, SEDAC could potentially have important roles in the areas of pursuing "technology transfer and commercialization opportunities" (Objective 5.4) and in fostering "new cooperative activities and partnerships among the NASA MTPE, university, and industry communities" (Objective 6.4).

More recently, NASA has attempted to improve its understanding of the EOSDIS user community. At the EOSDIS Potential User Group Conference held in Leesburg, Virginia in June 1995, MTPE brought together a wide range of "potential users" to characterize the needs and characteristics of the diverse user community including researchers, commercial and operational users, resource planners and managers, policy and decision makers, the legal community, educators, libraries, the press, and the general public. The preliminary draft conference report has been made available via the World Wide Web (WWW; see <http://bbq.ncgia.ucsb.edu:80/eosdis/newversion/title.html>) and in hard copy (NASA, 1995b). It is clear from this report that a wide range of users could benefit

from EOSDIS data and information resources, but that much remains to be done to make this happen.

In summary, it appears that SEDAC, including especially its Information Gateway efforts, can contribute to MTPE goals and strategies in a number of important ways. As detailed in this document, SEDAC can help make available a wide range of data and information that will help potential users from both the natural and social sciences take better advantage of EOSDIS data and information resources. Such efforts can help improve user awareness of other data and information services provided by EOSDIS and broaden and deepen the overall EOSDIS user community, e.g., by reaching out to communities not traditionally served by NASA or other remote sensing data providers. SEDAC can also help EOSDIS develop new channels for reaching the general public, the educational community, and other categories of users. Finally, specific SEDAC data and metadata holdings such as those related to population, land use, international environmental agreements, integrated assessment modeling, stratospheric ozone depletion, China, and Mexico may have direct application and utility in meeting specific MTPE research objectives.

2.0 The Nature of Global Change Research: Human Dimensions Context

This section discusses the interdisciplinary nature of the research on the human dimensions of global environmental change and the need to integrate the efforts of the natural and social sciences. The development of a predictive understanding of how human activities affect and are affected by the Earth system is a complex task. Human influences and interactions with the global environment cover spatial scales from local to global and extend over temporal scales ranging from minutes to centuries. All segments of human enterprise participate in this global experiment, both on individual and collective bases. As such, global change research must further focus on social and institutional systems as they relate to their physical, chemical, and biological counterparts. Effects on human society needs to be gauged, since each of us is intimately tied to the world about us.

The scientific and policy communities have increasingly recognized the importance of considering the human dimensions of global environmental change from an interdisciplinary perspective. These human dimensions include not only the mechanisms by which human activities may "drive" environmental change at a range of spatial and temporal scales, but also the myriad ways in which human behavior, perceptions, and responses are likely to interact with and affect evolving environmental changes or even the possibility of such changes. Understanding of both causes and effects is needed in considering mitigation and adaptation options regarding global environmental change.

Interactions between environmental and societal change can blur the distinction between the driving forces and the implications of global environmental change. For example, past and present environmental variability has already strongly influenced the dynamics of population growth and distribution and associated land use and will likely continue to do so in the future. Such linkages and feedback, which apply to almost every question of human behavior linked to the environment, confound attempts to separate "causes" from "effects."

As people in all walks of life consider the importance of environmental changes and what to do about them, they will need to know much more about their own role and contribution to such changes, the expected timing and distribution of potential changes, and the net effects of changes for themselves, their descendants, and other aspects of life that they value, given the alternative mitigative and adaptive actions available to them.

This broad challenge manifests itself in the research agendas of the U.S. Global Change Research Program (USGCRP) and other national and international research endeavors that address environmental change and its human dimensions. For example, the USGCRP is defined as encompassing "a program of focused studies to improve our

knowledge of the physical, geological, chemical, biological, and *social* processes that influence and govern Earth System behavior" (emphasis added). It specifically recognizes that:

Without an understanding of human interactions in global environmental change that is based both on empirical observations of human behavior and on a better understanding of the consequences of human actions, the models of physical and biological processes of change will be incomplete. Such links to human behavior are required for fully effective national and international policies to deal with global change (CEES, 1992).

The 1995 USGCRP plan (CENRR, 1995) emphasized the need for the development of a predictive understanding of how human activities are affecting and affected by the Earth system. To ensure a coordinated focus that addresses questions about global change and its human dimensions, the USGCRP is functionally organized into a framework of six coupled streams of research activity:

- *Observing the Earth System* includes (1) establishing an integrated, comprehensive, long-term program of land-, ocean-, in situ, and satellite-based observations on a global scale that monitor and describe the current state of the Earth system and (2) assembling and analyzing observations on recent and past social and environmental changes in the Earth system.
- *Managing Data and Information* includes assembling, processing, storing, and distributing data and information that document the state of the global system, both the conditions of the natural system and of the societal systems that are influencing and are influenced by global environmental changes.
- *Understanding Earth Processes* includes conducting a program of focused studies to measure, analyze, and investigate the physical, chemical, biological, geological processes and societal influences that govern Earth-system behavior and the interactions of human activities with the global environment.
- *Predicting Changes* includes developing, testing, and applying integrated conceptual and predictive models of the coupled Earth system in order to provide insights and projections of the response of the atmosphere, oceans, and land surface to natural and human influences and to reconcile predicted and observed Earth- system.
- *Analyzing Consequences* includes evaluating and interpreting the environmental and societal consequences and impacts of global change and understanding the potential for natural and technology-enhanced adaptation to and mitigation of global change.
- *Assessing Policies and Options* includes research on social and economic interactions, decision frameworks (especially those which include decision making under uncertainty), and development of the policy and economic tools of the various choices for responding to global change (CENRR, 1995: 22-23).

At the international level, the Human Dimensions of Global Environmental Change Programme (HDP) of the International Social Science Council (ISSC) has developed an overall framework for research on the human dimensions. It emphasizes that:

Research on this topic must consider human activities both as they contribute to and as they are affected by global environmental change. The human activities that interact with the Earth's natural systems are driven by three fundamental factors: the number of human beings and their distribution around the globe; human needs and desires, as conditioned by psychological, cultural, economic, and historical factors, which provide individuals and societies with motivations to act; and the cultural, social, economic, and political structures and institutions and norms and laws that shape and mediate their behavior (Jacobson and Price, 1990: 13).

It goes on to note that:

Research undertaken to serve this urgent global need will require conceptual, theoretical, and methodological development in the social sciences, both within disciplines and through the expansion of existing interdisciplinary fields of enquiry and the creation of new ones. Collaboration among social scientists and between social and natural scientists and cognizance of world-wide interconnections among the phenomena studied will be essential in order to deal with the crucial issues involved in global environmental change (Jacobson and Price, 1990: 14).

Efforts to improve understanding of the human dimensions of global environmental change thus pose complex challenges for the scientific community.

Much of the required research involves more than one scientific discipline and indeed demands integrated, interdisciplinary methods and data. Unlike traditional research in the *natural* sciences, human dimensions research must take into account the underlying values, perceptions, and behavioral choices that are unique to human beings and human societies. Unlike much traditional research in the *social* sciences which is site- or culture-specific, human dimensions research must address temporal and spatial scales, a need for quantitative modeling and prediction, and a degree of scientific collaboration that match what exists in the large, ongoing natural science programs to understand global environmental change. These challenges are especially apparent in the area of data and information related to the human dimensions, where existing data systems in the natural sciences are seldom matched by suitable counterparts in the social sciences and where wide disparities exist in scale, coverage, quality, and quantity of data.

3.0 Information Gateway Mission and Requirements

This section examines the functional elements of the Information Gateway and requirements associated with these functional elements. We begin by establishing a working definition of the Information Gateway and then explore three of its facets:

- 1) the requirements to provide access to socioeconomic data and information resources;
- 2) the need to provide access to natural science data and information resources; and
- 3) the need to support the integration of socioeconomic and natural science data and information.

The Information Gateway establishes access and service requirements to be rendered to all SEDAC user communities, both to the users defined in the Data and Applications Development Plan and to a wider range of users defined in this Information Gateway Plan. It is a "gateway" in the literal sense, a port of access to the EOSDIS and SEDAC data and information products. The functional requirements identified in this plan are translated into information services and system functions in subsequent SEDAC planning documents, particularly the Architecture and Operations Concept Document and the System Design Document.

The DADP addresses specific information products focused on needs of specific user communities. In addition to providing access to these application services, the Information Gateway Plan addresses a much broader community by providing that community access to information about the range of data and information products available in EOSDIS, and in other sectors dealing with relevant socioeconomic and environmental information. The creation of a data catalog interoperable with the GCMD, the basic tool used to document EOS products, is a primary additional contribution of the Gateway above and beyond the SEDAC applications products identified in the DADP. Combining the GCMD with a broader range of human and environmental metadata better enables interdisciplinary research and informed decision making by providing an integrated view of data and information products available. SEDAC proposes to accomplish the development of such a capability by enlisting the cooperation of other organizations. SEDAC will work closely with the CIESIN Information Cooperative to identify data resources and document those resources compatibly with the GCMD (see 4.1).

The SEDAC strategy in the development of the SEDAC catalog and metadata information system emphasizes access to a broad range of metadata maintained by the source organization. Unlike EOSDIS, which can employ a homogeneous approach to data documentation, SEDAC must address the realities of the management and delivery of metadata in light of heterogeneous information systems and a mix of standards. The

SEDAC Gateway system has been a successful attempt to achieve this, while remaining consistent with required metadata standards. The standards however evolve. It is understood that the Committee for Earth Observations Satellites (CEOS) is undertaking the definition of a common metadata protocol applicable the documentation and delivery of global change related remote sensing products of participating nation states. Development of the SEDAC Gateway system must remain aligned with the CEOS initiative while at the same time serving the needs of a broader and possibly less sophisticated user community, and serving information from providers who are not bound to CEOS protocols. This situation is a primary underlying reality facing the SEDAC Gateway system and poses a major challenge to the SEDAC system engineering effort.

This Plan articulate functional requirements relative to a matrix of seven service categories:

- Data and information resource catalog;
- Data and information resources at local and distributed archives;
- Data/information order and retrieval system;
- Information guides;
- Information resource discovery tools;
- Data analysis and processing tools; and
- User services support.

3.1 Definition

The Information Gateway is a two-way gateway between the socioeconomic and earth science data and information domains. The Information Gateway includes a directory capability, interoperable with the GCMD and EOSDIS, that provides the socioeconomic community with information about earth science products and the earth science research community with information about socioeconomic data. As a part of the Gateway effort, SEDAC also provides socioeconomic users and earth science users access to those data and information either acquired by SEDAC or made available by SEDAC through cooperative agreement with the source. The content of the archive is guided mainly by SEDAC applications development requirements.

The SEDAC Information Gateway efforts focus on recognized interdisciplinary areas of research that involve both socioeconomic and Earth Scientists in which needed data can be made accessible through SEDAC. Our understanding of research priorities is guided by the extensive literature developed through various consensus building efforts. These include efforts of the National Research Council (NRC) that address research issues in global environmental change; the considerable library of program definition efforts developed in documents issued by the USGCRP; and documents developed through similar efforts of the ISSC HDP.

3.2 User Community

The SEDAC DADP defines a target user community consisting of decision makers, analysts, and natural and social scientists who influence the decision making process related to a number of key global change issues. However, in practical terms, SEDAC applications development effort cannot hope to engage the interests of all possible users with the kinds of focused, value-added products and services it is developing. Only a subset of users with particular needs for these services are likely to utilize them. On the other hand, the Information Gateway effort provides the opportunity to serve a much larger proportion of this community with services designed to meet their broad interests in access to data and information. For example, although not all users will need specific tools that integrate Earth science and social science data, most will, in the course of their work, probably have use for more general types of services such as access to a catalog, encyclopedia, or dictionary. The Information Gateway aims to fill this need.

In attempting to meet the needs of the social and natural science user community engaged in interdisciplinary research on the environment and human interactions, SEDAC's Information Gateway can also meet the needs of a much broader user community. This broad user community includes educators and students interested in environmental management and human impacts, businesses and governmental bodies involved in environmental and resource management, and a variety of other user groups. SEDAC's ongoing User Model activities are attempting to better define these potential user communities, assess their characteristics and data and information needs, and evaluate possible channels for meeting these needs. For example, SEDAC has recently evaluated the needs of CGIAR data users and attempted to building a user-defined WWW interface to CGIAR metadata database.

Based on the recommendations of the EOSDIS Potential User Conference held in June 1995 (NASA, 1995b), SEDAC User Model activities will also focus on: 1) enhancing SEDAC's interaction with local, state, national, and international data and information systems; and 2) seeking innovative ways to educate the potential user community about the utility of its data, information, tools, and services.

3.3 Requirements for Access to Socioeconomic Data and Information Resources

Global change research has tended to focus on understanding how Earth science processes, such as climate patterns, are changing. Human activities within social systems, including population dynamics, economic systems, political systems, and technology development and implementation, have significant impact on the global environment. Understanding the human dimensions of global change—how global change affects human populations and how humans, in turn, affect the environment—requires that researchers, decision makers, and the world's peoples come to understand the combined impact on the environment of natural and social processes.

Understanding these complex interactions requires forging strong interdisciplinary links between the natural and social sciences, as well as providing needed information

to the decision makers of all nations who are confronting environmental and social issues. These activities must begin with sharing of information and knowledge that transcend political and cultural boundaries, providing accessible mechanisms that embrace users and providers of global change information throughout the world.

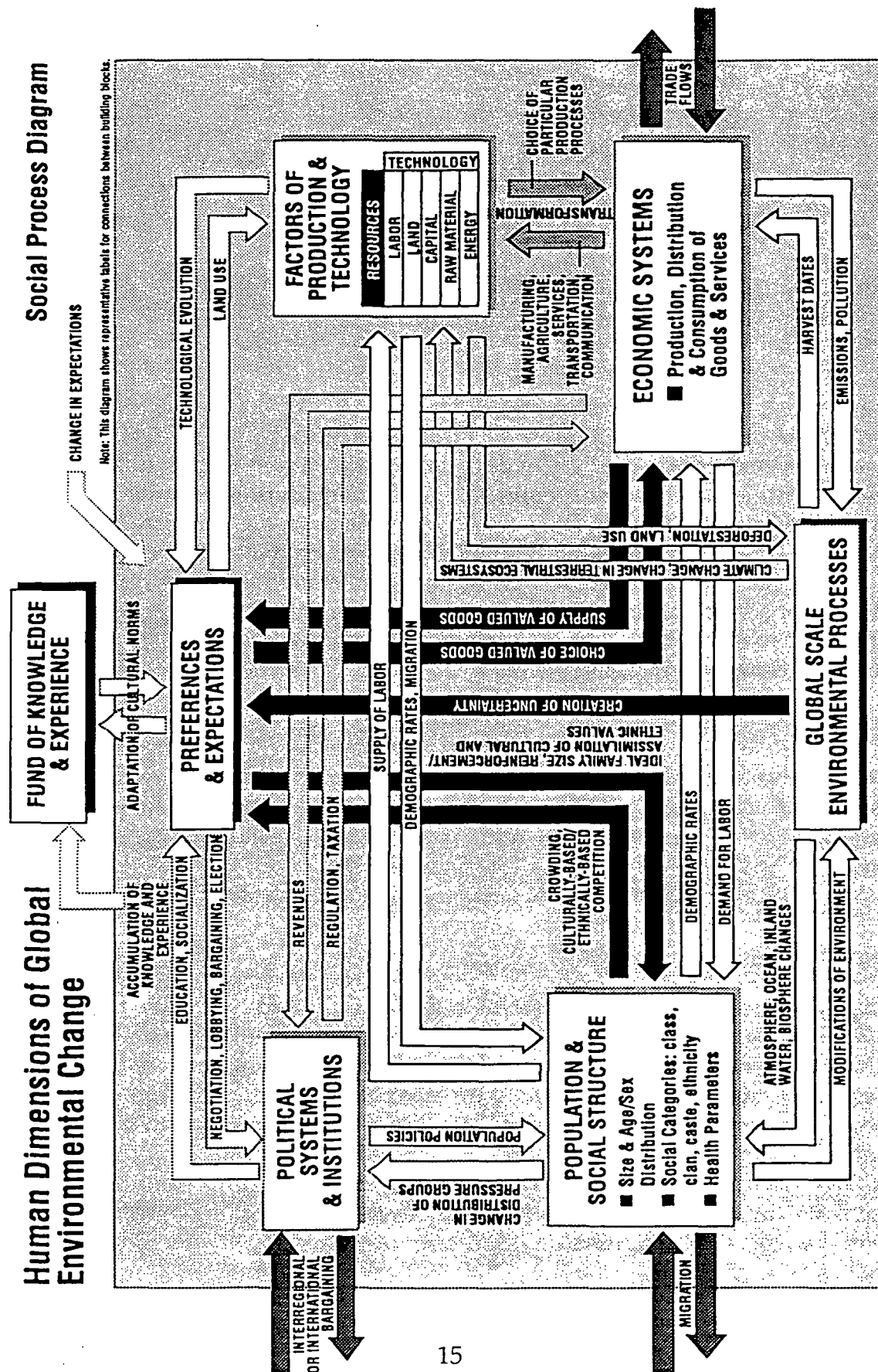
3.3.1 Science Requirements

The formulation of science requirements driving human dimensions research has continued over a number of years. The need for human science or socioeconomic data to understand the implications of global environmental change has been recognized by the USGCRP, as well as by NASA, the NRC, the HDP, and others. The HDP has produced a number of reports that attempt to focus human dimensions research on an international basis. National programs have significant human dimensions components, most notably the UK, Canadian and U.S. global change research programs. The HDP and the IGBP have joined in establishing a focused program of research on land use and land cover change, which has recently issued a draft research plan.

In 1991, CIESIN and the Aspen Global Change Institute (AGCI) commissioned a panel of social scientists to exchange ideas and examine how human social systems contribute to global change. The panel developed a model (Figure 1) showing how human activities and interactions affect and are affected by global change (CIESIN, 1992: 10). This model, called the Social Process Diagram, begins to establish the components of the human system that interact as forces contributing to Earth system processes. The model also establishes a useful foundation from which social science data and information requirements can be determined and a basis by which to guide SEDAC Information Gateway efforts. The model was designed to complement the Bretherton Diagram (discussed in the next subsection) which depicts the Earth System Model used in the natural sciences.

The Social Process Diagram is an important tool for scientists in several ways. First, it shows the relationships among the major social systems and how these affect and are affected by global change. It also indicates the directions and time scales of the various interactions among these systems and how these systems interact with natural systems. By showing these relationships, the role human institutions and processes play in global change can be illustrated. It also becomes easier to visualize and organize wide variety of studies that must fit together to address the implications of human activities in the environment. This includes asking the right questions, identifying the data and analysis tools to conduct interdisciplinary research, and using the appropriate integrating methodologies to model the interactive processes and address global change issues.

Figure 1. The Social Process Diagram (CIESIN, 1992).



The Social Process Diagram provides a basis for identifying eight principal science data domains used for the classification of data in SEDAC. These data domains are:

- Population Dynamics;
- Land and Freshwater Resources;
- Industry and Energy;
- Agriculture and Food Security;
- Economic Activity;
- Human Attitudes, Preferences, and Behavior;
- Policy and Institutions; and
- Human and Environmental Health.

As illustrated in Tables 1 and 2, these eight data categories define priority areas of interdisciplinary research on the human dimensions of global environmental change, as consistently identified in consensus reports of the NRC and in programmatic definitions of the USGCRP and the HDP. They also map well onto the Human Interactions Elements of the U.S. Natural Disaster Reduction Strategic Framework (CEES, 1992), and wider still, to generally recognized areas of data availability and the organization of scientists and of scientific research. The first four of the eight categories, and to some degree the next to the last and last, define areas in which social scientists and natural scientists share clearly overlapping interests.

Each of these data categories are updated below. The descriptions address, in part, the area of interdisciplinary research, the scientific disciplines involved in the research, and the associated scientific questions and unknowns that lend priority to the area. Specific priorities for the acquisition, archiving, and dissemination of specific datasets in each domain will be driven by the Data and Applications Development Plan and are specified in the SEDAC Science Data Plan. The IGP Update outlines areas within each domains that which will be the focus of data identification efforts during the next year.

Population Dynamics

Demographic data needed for purposes of global change studies include not only the total population and its change over time, but also key demographic parameters such as fertility, mortality, migration, and age structure at national and sub-national levels. Also of interest are other demographic variables such as housing, ethnicity, occupation, income, and access to services. These data are customarily recorded in various international statistical compendia drawn from national census or other national sources. In addition, household-level data are especially important for human dimensions studies because they provide insights into the underlying processes at work in human behavior in relationship to the environment (e.g., demographic and epidemiological transitions).

Table 1. Relation between SEDAC Science Data Domains and U.S. NRC reports concerning the human dimensions of global environmental change. Titles listed in the table reflect recommendations for research topics and research priorities in the case of NRC 1990, 1992, and 1994a and areas of discussion in the case of NRC 1994b.

Data Domain	NRC 1990	NRC 1992	NRC 1994a	NRC 1994b
Land and Freshwater Resources	Regional land cover conversion; Biomass burning	Land use and food production	Understanding land use change	Resource use and resource management
Industry and Energy	Global greenhouse gas emissions; Intensity of energy and materials use	Technological change; energy intensity; technology-environment relationship	Designing policy instruments and institutions to address energy-related environmental problems	Global economic development and environmental change
Agriculture and Food Security	Global agriculture	Land use and food production	Assessing impacts, vulnerability, and adaptation to global changes	Impacts and consequences of global environmental change; Resource use and resource management
Population Dynamics		Population growth	Understanding population dynamics and global change	Population dynamics and environmental change
Economic Activity		Economic growth		Global economic development and environmental change
Human Attitudes, Preferences, Behavior		Attitudes and beliefs; Valuing consequences of environmental change		Impacts and consequences of global environmental change
Policy and Institutions		Political-economic institutions; Decision making in response to global environmental change; Environmental conflict; International environmental cooperation	Improving policy analysis: research on the decision-making process; Designing policy instruments and institutions to address energy-related environmental problems	Policy science of global environmental change; Integrating knowledge about global environmental change
Human and Environmental Health				

Table 2. Relation between SEDAC Science Data Domains and USGCRP and HDP reports concerning the human dimensions of global environmental change. Titles listed in the table reflect recommendations for research topics and research priorities (CEES, 1993, 1994; CENRR, 1995; Jacobson and Price, 1990).

Data Domain	CEES 1993	CEES 1994	CENRR 1995	HDP 1990
Land and Freshwater Resources	Land use		Water quality and supply; Resource use & management	Land use; Social dimensions of resource use
Industry and Energy	Energy demands; Energy transformation		Reduction of greenhouse gases and ozone-depleting chemicals; Development of technologies to cope with climate change and increased UV radiation	Energy production and consumption; Industrial growth
Agriculture and Food Security		Response of agriculture to climate change scenarios	Effects of global change on agricultural systems	Environmental security and sustainable development
Population Dynamics	Population growth and distribution	Population growth as a cause or effect of global environmental change	International population trends	
Economic Activity	Economic behavior	Economic growth as a cause or effect of global environmental change; Benefits and costs of policy options	Trade and global economic activity; Costs of alternate technologies; Economic analyses and response strategies	
Human Attitudes, Preferences, Behavior		Impacts of global change on human activities	Awareness of and responses to changes in environmental conditions	Perception and assessment of global environmental conditions and change
Policy and Institutions	Legal and regulatory requirements	Effect of economic, political and legal systems on environmental responses; Needs of policy analysts; Processes through which policies are developed, implemented and evaluated		Impacts of local, national and international social, economic, and political structures and institutions
Human and Environmental Health		Impacts of global change on human health	Health impacts of environmental change	

Meeting the basic human needs of growing populations is dependent on a healthy environment. These human dimensions need to be given attention in developing comprehensive policies for sustainable development in the context of population growth. Demographic factors, combined with poverty and lack of access to resources in some areas, and excessive consumption and wasteful production patterns in others, cause or exacerbate problems of environmental degradation and resource depletion and thus inhibit sustainable development.

During the contract year, SEDAC will focus its data identification efforts on population distribution and migration especially in ecologically vulnerable ecosystems, interrelationships between population and resources, technology, economic, and urban growth, and human settlements.

Land and Freshwater Resources

Land-use and land-cover change is significant to a range of themes and issues central to the study of global environmental change. The alterations it effects in the surface of the earth hold major implications for sustainable development and livelihood systems and also contribute to changes in the biogeochemical cycles of the earth, affecting the atmospheric levels of greenhouse and other trace gases.

Land use and land cover changes are major driving forces of global environmental change, affecting the physical characteristics of the planet, global biogeochemical and hydrological cycles, availability and productivity of natural resources, and environmental quality. Socioeconomic variables that affect and are, in turn, affected by global land use change include population growth and distribution, affluence, and technology development and application. The value of land, in environmental terms, resides in the resources it provides, including the freshwater resources that are contained in rivers and lakes and subsurface storage.

"Land cover" describes the land surface of the Earth in terms of physical attributes in generalized categories such as "pasture," "cultivated land," or "deciduous forest." "Land use" provides information that describes the human and dynamic driving forces behind the observed pattern of land cover involving factors such as economics, property rights, and custom. It emphasizes human use of land and other terrestrial resources, especially where such use has led to degradation of economic productivity or other valued characteristics.

Principal land uses in global perspective include agricultural production, livestock grazing, forest product development, and managed open space. Less extensive, but significant uses include urbanization and transportation, preserves and recreation, and resource extraction. While the global use of land has changed continually through time, and particularly in the last century, it is expected that the pace and consequence of future land-use changes will increase dramatically in the next few decades, driven by the needs of an increasing world population. Understanding the driving forces behind such changes defines an important goal in research in global environmental change.

It seems also fair to say that in terms of future impacts on life, the quality of life, human health and biological diversity, the effects of future land use changes will far exceed any more direct effects of "climate change" or "greenhouse warming." Climate change and land use change are intimately related, and coupled as well with regional and global population change. Perhaps for these reasons, the effects of changing land use has been consistently identified as a priority area for study in both the USGCRP and the HDP. In 1995 the USGCRP has added emphases on water quality and supply and on resource use and management (see Table 2).

During the contract year, SEDAC will focus its data identification efforts on land use/cover, biomass burning practices, afforestation/reforestation efforts and maintenance of soil carbon pools, air/water quality, biodiversity, wetlands, infrastructure, and topography.

Industry and Energy

Data dealing with industrial activities and energy use enter the study of global environmental change principally through the efforts of researchers to trace the mass flows for key industrial materials of environmental significance. These include the waste emissions of carbon dioxide and other gaseous and solid emissions associated with industrial activities and the use and transformation of energy in various forms. It also encompasses the socioeconomic and technological forces that drive the evolution of industrial processes. These facets of industry and energy are sometimes studied in the sense of a metabolic process, in terms of industrial metabolism and industrial ecology.

Involved, in addition to chemical and physical processes, are also the dimensions of energy demands, industrial growth, consumption and the sustainable development of natural resources, technological change, technology-environment relationships, and ultimately the issues of the mitigation of and adaptation to global environmental change. As in the case of land use (or land and freshwater resources), the study of industry and energy has been ranked high in USGCRP and HDP priorities since the initial definition of these programs.

During the contract year, SEDAC will focus its data identification efforts on energy supply and fuel usage in the utility, industrial, commercial, and residential sectors; fuels and efficiency improvements within the transportation sector as well as transportation modes and influences of urban planning, and energy efficient technologies in industries such as chemicals, petroleum, cement, and pulp and paper.

Agriculture and Food Security

Agriculture and food security embrace a wide spectrum of information related to cropping systems, animal husbandry, agroforestry, biodiversity and genetic resources, fisheries management, dry land management, cultivation practices, soil and water management, irrigation systems, crop nutrient management, integrated pest management, post-harvest handling, agricultural products and processing, food and nutritional security, agricultural technology development and dissemination,

agricultural research methods, agricultural marketing, and agricultural and food policies.

Agricultural and natural resource management gained global recognition through documents such as the IUCN *World Conservation Strategy* and the Brundtland Commission's *Our Common Future*, and the United Nations Conference on Environment and Development (UNCED) of 1992. Data on human dimensions of agricultural and natural resource management provides insights on how global farming communities have interacted with their changing agricultural and natural resource environment including its floral and fauna resources. Recent studies show that human dimensions of agroecological zones, agricultural systems including agroforestry, and natural resource management are far more sophisticated and complex than previously assumed. Furthermore they offer new models for agricultural development that are both environmentally and socioculturally sound. On the other hand, in most instances, information on human dimensions of agriculture has never been systematically identified and documented, hence they may not exist or if they do they may not be easily accessible to users such as agricultural researchers, students, agricultural administrators, policy makers, extension specialists, and development practitioners.

Food production, the effects of global environmental change on agricultural systems and the response of agriculture to climate change scenarios have been cited as explicit research emphases in the USGCRP. Agriculture and food security are included implicitly in the HDP emphasis on environmental security and sustainable development.

During the contract year, SEDAC will focus its data identification efforts on fertilizer usage in the agriculture and forestry, and methane and nitrous oxide emissions.

Economic Activity

Studies of economic activity have been included in the program plan for the USGCRP since 1992. Areas for specific research emphasis have included economic growth and economic behavior; economic growth as a cause or effect of global environmental change; benefits and costs of policy options; trade and global economic activity; costs of alternate technologies; and economic analyses and response strategies. The HDP addressed economic data specifically in a 1992 HDP Report on Economic Data and the Human Dimensions of Global Environmental Change (Yohe and Segerson, 1992).

Formulation of strategies for long-term sustainable development require production and consumption data relating to all economic activities including industry, energy, agriculture, forestry, fisheries, transport, and infrastructure in order to optimize ecologically sound resource use and minimize waste. Macroeconomic and sectoral policies have, however, rarely given due attention to population considerations. Explicitly integrating population into economic and development strategies will both speed up the pace of sustainable development and poverty alleviation and contribute to the achievement of population objectives and an improved quality of life of the population.

During the contract year, SEDAC will focus its data identification efforts on impact of international trade patterns on the use and value of environmental resources such as wetlands, coastal zones, and forests; effects of economic liberalization on the use of environmental goods,

Human Attitudes, Preferences and Behavior

Reduced to simplest terms, the most basic units of the human dimensions of anything are the actions of individuals, which can be understood in terms of personal attitudes, personal preferences, and ultimately, personal behavior. Although slippery and to some degree inherently unpredictable, these human quantities are undeniably fundamentally entwined in the study of global environmental change.

Human attitudes and behavior including valuation of the environment, issues of equity among generations, global change decision making, and the role of institutions becomes an essential element for assessing the human dimension of global change. Without these foundations, the strategies for mitigation or adaptation may be misguided, resulting in either inadequate or unnecessary policies and programs.

The USGCRP has identified human attitudes and beliefs, valuing consequences of environmental change, the impacts of global change on human activities, and awareness of and responses to changes in environmental conditions in its definitions of priority research areas for the program. The HDP emphasis has been on the perception and assessment of global environmental conditions and change.

Datasets for the application of human attitudes, preference, and behavior come from disciplines such as anthropology, sociology, history, and psychology; such trends can be tracked, in real time, to some degree by survey techniques, which are among the principal tools of the social sciences.

Policy and Institutions

Institutions and organizations, ranging from families to governments, influence policy formation, the way society is organized, and ultimately and in most cases, the ways that individuals react to global environmental change. Knowledge of these often-changing structures are thus intimately linked to human response and behavior.

Of particular relevance is the large body of data regarding the development and implementation of international, national, and sub-national agreements concerning global environmental change, including activity in the executive, legislative, and judicial branches of government as well as contributions from a variety of non-governmental actors.

The USGCRP has identified a large number of research emphases for the study of the human dimensions of global environmental change that help define this data category more explicitly. These include political-economic institutions; international environmental cooperation; decision making in response to global environmental change; environmental conflict; legal and regulatory requirements regarding energy

use; effect of economic, political and legal systems on environmental responses; needs of policy analysts; and processes through which policies are developed, implemented and evaluated. The relevant HDP emphasis is that of impacts of local, national and international social, economic, and political structures and institutions.

During the contract year, SEDAC will focus its data identification efforts on national-level sustainable development indicators.

Human and Environmental Health

This category includes data holdings on human health related to the environment, ranging from pollution data to epidemiology. Also of interest are data on the health of the environment itself as it may relate to human health. An obvious area of interest that couples physical, chemical and human data are aspects of the epidemiology of skin cancer, including malignant melanoma, as these are related to the transparency of the atmosphere and the varying ultraviolet radiation from the Sun.

The USGCRP included explicit emphasis on research in human and environmental health in its FY 1994 program plan, and again in FY 1995, calling explicitly for studies of impacts of global change on human health and on the health impacts of environmental change, including studies on CFC chemical replacements and UV. This reflects in part the entry of the U.S. Department of Health and Human Services and the National Institutes of Health as agency participants in the program.

During the contract year, SEDAC will focus its data identification efforts on births, mortality, morbidity, impacts of UV exposure on the immune system, aging process, and sensitive tissues; health effects of pollutants; toxicity of pesticides; and occupational exposure to health hazards.

3.3.2. Data Identification

Unlike most of the other EOSDIS data centers, SEDAC does not have direct responsibility for managing the data stream from an in-orbit EOS sensor. Nor can it immediately draw on a well-defined scientific community that has a strong tradition of collaboration and consensus-building in defining the parameters that are to be measured by and archived from EOS payload instruments. Instead, SEDAC must identify and then develop a working relationship with a diverse set of data sources, ranging from the individual research scientist concerned with a relevant case study to large governmental agencies that collect and analyze data on national or international economic activity. This diversity of sources makes it imperative that SEDAC work aggressively to identify and access important extant datasets, to determine where there may be critical gaps, and to provide leadership in filling in these gaps through efforts at the directory level and at the data access level. It must also work closely with the relevant scientific and data communities to build a sense of identity and cooperation and to support the establishment of mechanisms for effective and efficient acquisition, processing, archiving, and dissemination of highly valuable data and information resources.

During the contract year, SEDAC will identify and describe key socioeconomic datasets related to areas mentioned in the previous section using the following strategies:

- Conduct inventories of data projects within each country;
- Identify clearinghouses of metadata, publications, and updates to past datasets;
- Browse World Wide Web periodically;
- Maintain working relationships with data providers and research organizations; and
- Participate in conferences and other outreach activities.

3.3.3. Information Gateway Services

This section examines Information Gateway services SEDAC has deployed to support user access, including planned enhancements to these services in the current contract year.

Data and Information Resource Catalog

When researchers extend their studies into disciplines outside their own, they require access to that discipline's data. As the NRC points out:

Because research on global change is so broad and draws on data, empirical research, and analysis from a wide variety of disciplines and countries, few social or natural scientists are familiar with more than a narrow portion of it. Giving them access to the data they need presents a major task for managing, storing, and distributing scientific information (NRC, 1992: 206).

In many cases, useful quality data exists, but:

...the fact that good information exists *somewhere* is of little help when one does not know where to find it (Croze and Vandeweerd, 1991: 121).

To overcome this barrier to interdisciplinary research, researchers need access to a catalog of extant data and information resources.

The current state of socioeconomic data catalogs is parallel to environmental data catalogs. First, metadata (data about data) may not exist, or when it does, is in unfamiliar or inconsistent form. Existing metadata are often poorly organized and maintained further limiting or confounding access to the data. To overcome these barriers, metadata will need to be developed and made available through a catalog system that documents a wide variety of socioeconomic data. It is furthermore critical that the methodology employed is fully compatible with environmental data catalogs, such as the GCMD, currently in development. The metadata and catalog system will

also need to protect the user from information overload by providing an easy to use interface, powerful and flexible searching tools, and an efficient metadata data model.

Another aspect associated with locating socioeconomic data and metadata is that they are archived at many separate organizations.

Not all significant public-domain data are in national repositories. This situation has developed for a variety of reasons: economic policy, research efficiency, and the personalities of principal investigators. Moreover, the secondary user is often unaware of the range of data available through the data centers and maybe unaware even of the centers' existence (NRC, 1991: 24).

This is not unlike the situation in environmental domain, but is made further complex by nature of the fact that there is no uniform instrument (such as a satellite) by which the data are originally collected. There needs to be a master catalog that contains metadata for many data centers. However, rather than creating a single, central catalog, each center should be responsible for maintaining its own metadata. A 1992 NRC report recommends:

An information network should be designed for global change research (including the human dimensions) that will make catalogues, data bases, data transformation, documentation, scientific literature, and other public information more widely and inexpensively available (NRC, 1992: 210).

This distributed approach would allow each data center to maintain their metadata, while at the same time presenting a single access point for the user.

The simplest approach to creating a distributed system would be to implement a standard solution that would be used by all data centers. This approach, though, would meet with considerable resistance. As Croze and Vandeweerd (1991: 121) point out:

There exists an abundance of information systems, large and small, public and private. It will never be possible to get everyone in the world to use one standard system, nor would this be wise, given the necessity for sectoral experts to continue contributing to information-generation within their own sectors.

If a standard system is not possible, the catalog must be able to be interoperable with heterogeneous systems, providing a glue to connect multiple hardware, software, and metadata systems and standards. Section 5.2, the Information Cooperative, describes the methodology by which holders of key socioeconomic data can cooperate in coordinating cataloging needs.

In the first contract year, SEDAC addressed the needs described above by developing an operational Data and Information Resource catalog. Dubbed "Gateway," the system traverses a distributed collection of heterogeneous resources to provide intuitive access

to socioeconomic data and information. The catalog serves to accomplish the following functions:

- Document extant socioeconomic data with directory metadata;
- Reflect extant data resources according to domains previously established in the IGP;
- When available, provide inventory and guide information;
- Facilitate multidisciplinary and interdisciplinary access to metadata and data;
- Conform to metadata standards as designated and required by EOSDIS;
- Incorporate an indexing vocabulary which is appropriate for describing the metadata content and which will facilitate metadata retrieval by a diversity of user groups;
- Accessible via an electronic, distributed catalog; and
- Flexible display to accommodate a diversity of user needs, e.g.:
 - Directory view* - Permits the user to discover an appropriate dataset or collection of datasets;
 - Inventory view* - Enables the user to determine details of how to obtain a specific dataset;
 - Guide view(s)* - Supplies supporting and supplementary documentation that allows the user to effectively use the dataset.

During this contract year, SEDAC will enhance the Data and Information Resource catalog based on lessons learned from the operational system, recommendations of the SEDAC UWG, and ongoing requirements analyses. Furthermore, the catalog will be enhanced by establishing additional "nodes" at high priority Information Cooperative sites. SEDAC will provide the technical infrastructure by which new nodes can easily be brought online and provide directory-level metadata. This will include software and user support to allow member sites to host "Gateway servers" which become a part of the SEDAC distributed archive and are accessible via the catalog.

Archive of Data and Information Resources

Although a researcher has located the source of a dataset, the data may still not be accessible for a variety of reasons, including lack of network connectivity, high cost, software incompatibilities, and poor documentation. As with metadata, data and information are held in many physical and institutional locations. To overcome these barriers, data and information resources must be linked together in a virtual archive. Each data center maintains its own data archive, preferably online. Also, because some data centers do not have the capabilities for providing critical datasets online, SEDAC is prepared to serve as an archive of last resort. The local SEDAC archive serves as another node in the larger virtual archive and guides its collection on the basis of DADP requirements. Section 5.3 outlines the role of the local SEDAC archive in supporting the

needs of the SEDAC user community. Primary contents and functionality of the archive are as follows:

- Data and Applications Development Plan information products developed from NASA Earth science data and socioeconomic data;
- General socioeconomic datasets, bibliographic databases, and information resources;
- Access to dataset and model guide information;
- Online automated access to most datasets;
- Availability of data on required EOS media; and
- Life cycle management capabilities for ingest, quality and integrity checking, status tracking, access restrictions, reporting, and file management.

Order and Retrieval System

To allow access to the SEDAC virtual archive an online order and retrieval system is required. This system should be part of the same system that provides access to metadata and utilize a similar interface. Online retrieval should be an important aspect of this system. This reduces the time delays and costs associated with copying data to magnetic media and shipping the magnetic media to the researcher. Also, each node in the virtual archive may have different and variable requirements for licensing and costs, and the ordering system will need to provide capabilities for supporting this type of flexibility.

The SEDAC Data and Information Resource catalog includes an order and retrieval system which addresses the following functional requirements:

- Provide distributed online catalog directory and inventory metadata access to data held at each site;
- Provide online display of dataset guide information, general access information, use restrictions, terms of agreements, licensing agreements, and differential access and pricing, based on users' online assertions and responses;
- Enable users to fill out an online order form and licensing agreement (if required) and capability for conducting user authentication (checking of user inputs against preset criteria);
- Determine class of user and implement appropriate information display, pricing, access, depending on class of user;
- Provide for automated sending and receiving of order forms and licensing validations by facsimile and or electronic mail;
- Provide for encrypted online authentication of users for access to data that is confidential and/or which is restricted to very limited audiences;

- Provide user support, accounting, report generation, status tracking, and data and information life cycle management capabilities as integrated components of the order system;
- Provide digital data storage space for online and offline holdings;
- Provide automated capabilities for online access (direct download and FTP) of frequently used data and information;
- Provide efficient systems for developing, loading, and modifying the directory, guide, inventory metadata related ordering information (authentication criteria, terms of agreements, etc.), status tracking, and data management information;
- Provide a common look and feel and common underlying protocols, across the order system interfaces for all nodes, for all of the above capabilities;
- Provide capability for dissemination by traditional media, as well as online; and
- Provide capability for customized processing and analysis services.

Information Guides

Facilitating the searching for, identification of, and access to data helps the researcher get the data, but usage may be limited because of difficulties working with the data. This problem can be dealt with by expanding metadata to include quality control information and to provide guides to the data that supply important information to facilitate usage of the data. Additionally, information will need to be provided related to each of the nodes in the virtual archive. This can be accomplished with an organizational guide. These capabilities can also be supplemented with thematic guides that provide expert reviews of a data domain or global change relevant issue along with online access to key literature in the area. The guide is a fundamental element of the metadata information model described in Section 5.1.

In the first contract year, SEDAC has developed Information guide capabilities to meet the following requirements:

- Dataset guides: provide detail documentation for the intent of datasets referenced at the directory level.
- Organization guides: describe cooperating institutions.
- Thematic guides: including science guides, regional guides, product guide document information from an interdisciplinary perspective.

The Information guides are linked with the Data and Information Resource catalog to provide an integrated mechanism for information search, browse, and order. During this contract year, these linkages will be enhanced.

Enhanced Information Resource Discovery Tools

SEDAC has released an initial version of the Gateway based on the Client/Server paradigm. SEDAC recognized the need to respond to the recent explosive growth of the

Internet, fueled primarily by the development of the World Wide Web and at the same time recognize that the World Wide Web is not yet universally acceptable. The user-friendly, point-and-click approach of Web "browsers" has made the Internet more accessible than ever before. SEDAC already uses the World Wide Web to provide access to Information Guides, and will develop additional Web-based applications to support a broader range of functionality. SEDAC will address needs of users' without WWW access through support of an e-mail interface. These augmentations of Gateway are discussed in the following:

SEDAC will leverage off the popularity of the World Wide Web by implementing Web-based metadata search, browse, and delivery tools to supplement the functionality of SEDAC's Data and Information Resource catalog (i.e., "Gateway" system). Access to selected SEDAC holdings will also be available via standard Web browsers. An enhancement to this service will allow data providers to register metadata via the World Wide Web (subject, of course, to quality control and relevance assurance procedures), enabling SEDAC to provide a more comprehensive socioeconomic and environmental resource catalog and "closing the loop" with SEDAC users by enabling them to participate more actively in developing the content of the Gateway resource.

The Web Gateway tool will support the needs of SEDAC users who use a Web browser as a standard means of information access and for whatever reason do not have a Gateway client installed. The number of such additional SEDAC users is potentially very large, since the total number of users who utilize Web browsers is in the millions and growing exponentially. This means that a dramatically increased SEDAC user base will now be able to take full advantage of the Gateway's distributed content. Increased usage will in turn stimulate data provider interest in setting up servers to participate in the Gateway network. For these reasons, SEDAC believes that development of a Web-based Gateway client is the single most cost-effective step that SEDAC can take to stimulate and improve access to the SEDAC Gateway. Accordingly, this will be the highest priority Information Gateway Plan activity in the current contract year.

A WWW interface to the Gateway would be an important capability for SEDAC's Information Gateway activities by increasing the visibility and ease of access to SEDAC's metadata holdings to a wide audience of WWW users. This approach would lower the barrier of users being required to install specialized Gateway clients in order to use Gateway, and would also broaden access to users with platforms for which clients do not exist such as the Apple Macintosh. The Gateway functionality currently provided by the MS-Windows and UNIX X-Windows clients could be provided using a set of WWW forms and pages and bridging software to pass queries from WWW to the Gateway and return results from the Gateway to WWW.

An additional consideration is that timeliness is critical in determining a service's niche in the emerging Web-based "content architecture of the Internet." Consider the dramatic array of new services that have entered use and become recognized as near-standards in recent months--e.g., Lycos and Yahoo. SEDAC believes there is an opportunity for the Web Gateway tool to become similarly recognized as a global WWW resource that is the best source for access to quality-controlled, focused socioeconomic information.

SEDAC will serve the needs of users who have e-mail-only access to the Internet in two ways: by providing an e-mail-based Gateway client and by assuring its users of access to a reliable e-mail to Web interface. The W3 Consortium has recently indicated that it is looking for mirror sites to share the system load of providing its Agora e-mail to Web service, which indicates, first, that it may not be sufficient for SEDAC to rely upon external services to provide e-mail to Web connectivity for its e-mail-only users, and second, that there may be an opportunity for SEDAC to provide a useful service, if appropriately structured.

SEDAC has received permission from the W3 Consortium to install a copy of its Agora e-mail to Web server for evaluation purposes. The Agora system does not currently support forms, although its developers indicate that they are planning to add this functionality in the very near future. For this reason, SEDAC will carefully evaluate the technical merits of Agora and other mail to Web solutions and will evaluate the most cost effective way to provide full Web form functionality via e-mail. SEDAC will investigate the possibility of providing such services to the global change community in a way that does not produce undue system load but still meets the objective of providing SEDAC users with reliable e-mail access to SEDAC-relevant Web content. It is important to note that SEDAC-relevant web content is not limited to Web content online at SEDAC or CIESIN, but rather is distributed worldwide.

Data Processing and Analysis Capabilities

Users of socioeconomic data will need to utilize a variety of tools for data processing and analysis. Generally, these tools, such as statistical analysis packages, will be provided and used within the user's local computing environment. However, certain functionality will need to be offered through SEDAC for some datasets:

- 1) *Browsing*. Browsing allows a user to explore the data by viewing static browse products (pre-generated examples of the data) or by using interactive browse tools that allow the user to select what views of the data to see. Browsing tools and products do not require the full functionality of statistical processing or visualization systems required by researchers, and they can be utilized by applied users to discover the information they require without a great deal of technical knowledge. Static browse products can be image or textual data and should be displayed as part of the catalog system. Dynamic browse tools should allow interactive exploratory analysis of data.
- 2) *Extractions*. Extraction capabilities allow users to define and generate a subset of the data. This subset can then be used by the user in a spreadsheet, charting package, or even a statistical analysis package for further analysis. Extractions are created on demand and would be available to the user after a variable period of time, depending on the size of the dataset. In general, extracts would be available in a fashion consistent with other data products.

- 3) *Visualization*. Visualization includes a range of tools and approaches for creating maps, charts, or other pictorial representations of data. Visualization is incredible powerful for applied users because it can easily and efficient communicate information. Visualization is often combined with browsing capabilities. Visualization requirements will be established by the DADP.

Information Gateway Web Page

During the contract year, SEDAC will continue to maintain the SEDAC Information Gateway page. The Gateway page will provide access to current research related to SEDAC social science domains, topical indices within the domains, abstracts of key datasets, and thematic guides.

Supporting User Services and Training

During the contract year, User Services supports user access to socioeconomic data by fulfilling the following requirements:

- Provide a help desk;
- Assist users to identify and locate datasets;
- Coordinate with other centers to assist users;
- Track data order accuracy and completion;
- Compile user service statistics;
- Provide system demonstrations;
- Staff user support office with science and data knowledgeable personnel;
- Update the bulletin board with the latest news;
- Provide 40 hour prime time staffing;
- Knowledgeable of processing history and product revisions;
- Provide up-to-date validation information; and
- Collect and compile data quality comments from users.

Additionally, User Support personnel will:

- Respond to user needs providing data, information, training, products, literature, and services;
- Prepare user contact information as part of the overall monthly report;
- Organize user outreach activities and workshops as appropriate; and
- Demonstrate how new products and services satisfy user requirements.

3.4 Requirements for Access to Physical Science Resources

Global change research encompasses a wide range of studies in the physical and life sciences. Such efforts have led to a rapid expansion of observational programs at national and international levels and a vast increase in the quantity and complexity of available data and information on natural systems. Remote sensing data in particular pose a significant challenge in terms of managing and disseminating large datasets to users in the science community. This challenge is especially evident in the social sciences where experience in handling the types and volumes of data from remote sensing platforms is more limited than in the natural sciences.

Global change research intrinsically demand the integration of diverse data in order to address both the natural and social science aspects of the particular issue. With advances in remote sensing technology, the amount of atmospheric and geophysical data has increased exponentially over the past several years. Easy access to the Internet and improved electronic communications have made such data much more accessible as well.

Another difficulty arising from this seemingly advantageous scenario of plentiful data sets and sources to choose from is sorting through the myriad of sources and data contents to find the specific information required for the application. This situation becomes amplified when users outside a particular discipline attempt to find and use data from that field of study for interdisciplinary applications. One of the goals of the Information Gateway is to make available critical physical and natural science data sets to the social science community that can be employed for interdisciplinary activities. Providing access to such data that have been deemed important to identified user communities through the Information Gateway is an important resource to the global change effort as a whole.

3.4.1 Science Requirements

In the mid-1980s, a group of natural scientists assembled at Jackson Hole, Wyoming, to construct a diagram showing the interactions among natural processes that influence global change. This diagram (Figure 2) is known as the Bretherton Diagram.²

The group that developed the Bretherton Diagram realized that understanding the relationships among their disciplines was crucial to studying the natural aspects of global change. The two major components of the Diagram, the Physical Climate System and Biochemical Processes, have several subcomponents. The Bretherton Diagram shows that information and data from one system or discipline is needed to evaluate processes in another system.

² Named after Dr. Francis Bretherton, who participated in this effort and chaired the Earth System Sciences Committee of the NASA Advisory Council.

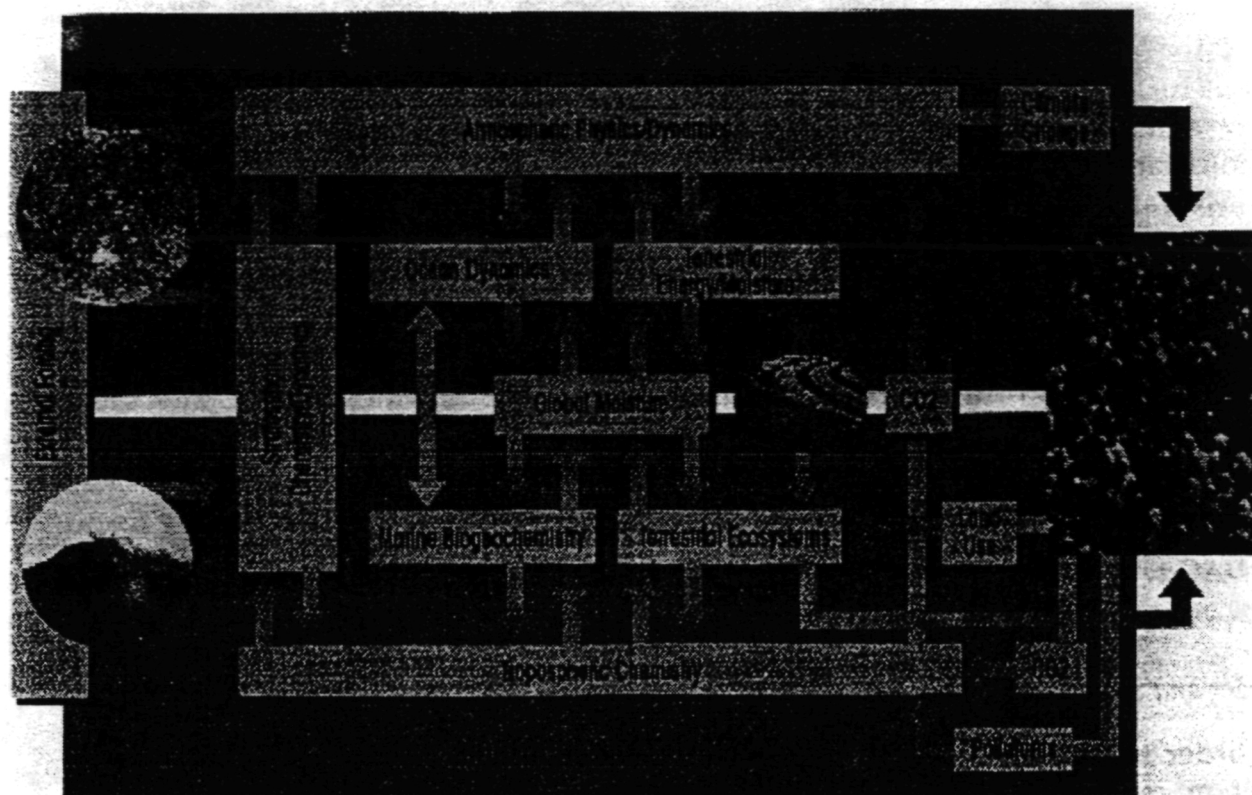


Figure 2. Simplified version of the "Bretherton" diagram. This diagram illustrates linkages between the physical climate system, biogeochemical cycles, and human activities (NSF, 1988).

Although "human activities" are noted on the Bretherton Diagram, the relationships within that category were not developed; the Social Process Diagram was an attempt to address the interrelationships in a fashion similar to the processes within the natural science system. Thus, the two diagrams are complementary, completing a conceptual model of the components and processes affecting and affected by global change.

These two diagrams also fundamental to understanding the concepts underlying the establishment of EOSDIS. The EOSDIS will manage data resulting from NASA's Earth science research satellites, field measurement programs, and other data essential for the interpretation of these measurements. The nine data centers (including SEDAC, which addresses socioeconomic data requirements) were established to manage the natural and social science data suggested by the two diagrams.

3.4.2 Information Gateway Services

In this section we identify SEDAC services to support access to information about physical science data.

Data and Information Resource Catalog

Social scientists seeking to use physical science data are confronted with similar problems as physical scientists searching for socioeconomic data. Hence, the needs are very similar. A catalog of metadata needs to be available for searching for and identification of appropriate data resources. Efforts such as the NOAA directory, the GCMD, and EOSDIS provide the basis for solving this problem, but social scientists need to be aware of the existence of these directories and have access to them. The ideal solution would be to make the catalog system for socioeconomic data interoperable with the physical science metadata, thus providing a tool that would allow researchers of all disciplines that would bridge both categories of data. SEDAC users will be provided access to physical science catalog systems including the GCDIS Global Change Master Directory, and EOSDIS IMS.

Local and Distributed Archive of Data and Information Resources

SEDAC will also provide interoperability to the other EOSDIS archives, thus providing a gateway to a primary collection of relevant physical science data.

Order and Retrieval System

In order to provide access to the EOSDIS archives, the SEDAC order and retrieval system must have interoperability with the EOSDIS order and retrieval system, allowing a user to submit a request through the SEDAC to another DAAC.

Information Guides

Social scientists will also need access to additional metadata that provides necessary contextual information. During the contract year, SEDAC will take measures to provide interoperability between SEDAC guides and the guides developed by other DAAC's in EOSDIS. SEDAC will supplement this with guides designed particularly for social scientists.

Data Processing and Analysis Capabilities

SEDAC will utilize the data processing and analysis capabilities of the larger EOSDIS system for processing physical science data and information resources.

Supporting User Services and Training

SEDAC user services will need to support the social science community in accessing EOSDIS data products. To accomplish this User Services will:

- assist users in using the EOSDIS V(0) systems;
- develop and maintain a dictionary of DAAC unique services; and
- remain cognizant of EOS capabilities and EOSDIS data products.

3.5 Requirements for Integration of Physical and Socioeconomic Data

The causes and consequences of human activities need to be factored into the discussions, research, predictions, policies, and actions surrounding global change. When the links between the natural and human domains are identified and quantified, we will begin to understand more fully the direction of global change. Achieving this goal requires that social and natural scientists work together to identify and understand global change and to recommend courses of action to ensure the long-term well-being of the Earth and all its inhabitants. The need to integrate socioeconomic and environmental data to fully understand Earth as a system, to guide the development of policy to mitigate effects, and to model the system's adaptive capacity requires SEDAC to provide certain basic services directly related to the work specified in the DADP.

The need for integration of environmental and socioeconomic data has been identified consistently in a wide range of national and international documents. For example, the recent National Research Council report on *Science Priorities for the Human Dimensions of Global Change* notes:

The [USGCRP] program expansion challenges the scientific community and the agencies that implement it to develop knowledge simultaneously in several major substantive and methodological areas that are new to the program. The challenge is increased by the fact that each of these new areas depends on knowledge about the human dimensions of global change, a relatively underdeveloped sector of the USGCRP. The expansion also requires the integration of social science research with the natural science research that has until now formed the core of the U.S. program. This kind of integration has been difficult in the past because of major structural barriers both in academia and government (NRC, 1994:3-4).

3.5.1 Science Requirements

Anthropogenic influences on Earth system processes necessitate examining the Earth system as a biogeochemical process with a human interaction and feedback system. Moreover, human response through legislation (mitigation) and behavioral change (adaptation) interact with the process. That is, the Earth system responds to inputs from the human system and in turn the human system responds to environmental demands. The human system has always been dependent on the environmental system, human pressure now binds the fate of the environmental system to it. Hence, understanding and, more importantly, predicting the behavior of the environmental system, demands not just a measurement of environmental phenomenology, but the examination of human influences. The need to understand the complex interactions between human and physical systems results in the need to integrate measured phenomenon. The integration of data and information will have explicit and implicit manifestations.

Geographic Information Systems enables integrated analysis of information layers. To achieve this, data must be associated spatially. Generally environmental measurements are attributed with a spatial location. Remotely sensed data are implicitly spatial.

Environmental measurements, such as water quality, are identifiable to a specific latitude, longitude and altitude. Though social science data can be spatial in nature, other aspects of the measurements, such as social structure, are usually the primary "locational" attributes. When spatial attribution is employed, it is often a broad location such as an administrative district. Hence spatial scale can be inconsistent when integrating social science and environmental data. This requires care when performing geographic analysis of joint features. Explicit association of environmental and socioeconomic data will require appropriate tools and interpolation models to enable integration on a spatial basis.

The integration of climate models with social process models reveals a second approach to integrate environmental and socioeconomic data that relies more on implicit association of data. In this case, modelers may estimate parameters driving the climate equation from socioeconomic data without explicitly associating the data on a variable by variable basis. The issue of scale is less important, but the need to establish comparability of the data underlines the importance of fully understanding the collection methodology and assumptions of the parameters being used. For example, worldwide population projections may be derived from varying models, some of which may consider environmental factors, while others do not. Depending on the assumptions made in the projection, the data may or may not be valid for integration in a climate change model that is driven in part by population pressure.

The problems of spatial scale, attribution, measurement methodology and assumptions are all crucial in providing scientists means to support the integration of environmental and socioeconomic data. SEDAC Information Gateway services will begin to address mechanisms needed to support the data integration needs of the SEDAC user community.

3.5.2 Information Gateway Services

In this section we identify SEDAC services to support access to information about integrated socioeconomic and physical data.

Data and Information Resource Catalog

During the contract year, SEDAC will continue to catalog integrated data products using the procedures described previously.

Local and Distributed Archive of Data and Information Resources

SEDAC will need to store integrated data products developed as a result of the DADP as part of its archive.

Order and Retrieval System

In addition the base level of functionality described earlier, the order and retrieval system will also need to include capabilities for ordering integrated data products and

provide those products in a common data distribution format such as the hierarchical data format to facilitate data integration.

Information Guides

During the contract year, SEDAC will continue to produce to dataset guides, model guides, and organizational guides. The requirements for information guides will be established the Science Data Plan Update.

Data Processing and Analysis Capabilities

Data processing and analysis play an essential role with integrated data. In addition to the functionality described earlier, a new capability must be included as part of the SEDAC data processing and analysis functions:

- **Ulysses:** During the past year, SEDAC has supported the development of Ulysses, an interactive tabulation/analysis system that allows users to generate cross-tabulations from the one-percent 1980 and 1990 PUMS files. During the contract year, SEDAC will continue to update the Ulysses by adding one-percent 1970 PUMS files. SEDAC User Services will also help users by helping them in extracting the data for more intensive analysis.
- **Model Visualization and Analysis Service:** SEDAC has developed the Model Visualization and Analysis Service to view, explore, and work with the data from precomputed model runs for the integrated assessment models documented in the model guides. During the contract year, SEDAC will continue to enhance the Service that enables users to perform exploratory analysis and visualization of archived IAM data.

Maintaining SEDAC Home Page

During the contract year, SEDAC will continue to maintain its home page. This home page will continue to provide access to SEDAC's application products and services.

Supporting User Services and Training

In addition to standard users services and training functions, electronic communication services will need to be supported that facilitate communication between different disciplines. To accomplish this, SEDAC will continue to support the use of the Human Dimensions Kiosk which enables sharing of research results among scientists.

4.0 Information Gateway Implementation Mechanisms

This section discusses the mechanisms necessary to accomplish the development and delivery of SEDAC Information Gateway products and services.

4.1 Information Model

The SEDAC Information System is implemented using an integrated, extensible information model as its underlying conceptual framework. This framework supports multiple views of metadata content, allowing support of a diversity of metadata standards and user groups. This section describes the information model and defines the protocols and standards which are employed in the creation and development of the SEDAC Information System.

4.1.1 Metadata Content

Metadata coupled with electronic search and retrieval tools become an information system which uses metadata as the primary integrating mechanism. This provides users with access to geographically distributed data resources, from multiple disciplines. SEDAC recognizes that different types of data resources from different disciplines are best represented using the metadata content most appropriate for each. That is to say, bibliographic datasets are best described using the MACHine Readable Cataloging (MARC) standard; spatially referenced datasets using the Content Standards for Digital Geospatial Metadata; satellite data using the Directory Interchange Format (DIF); and so forth. SEDAC works to define the particular metadata content needs of socioeconomic data resources.

While commonalties exist across metadata content standards (e.g., title, abstract, keyword), there are also metadata content requirements that suit the needs of a select community (e.g., sensor, latitude, collection method). Designing a metadata schema that allows for the particular needs of each of many data providers adds value to the information system. In such a model, the onus of integrating across, and translating between, different metadata characterizations must fall on the information technology that provides the directory service. In support of this philosophical approach, SEDAC's metadata model allows metadata of multiple formats to be integrated through the SEDAC Gateway.

4.1.2 Information Model: Logical Organization

The Committee on Earth Observations Satellites (CEOS) Model is a metadata framework suggested for use by the international community for describing data from satellite sources. SEDAC has adopted and extended the CEOS model and plans to

continue developing within its framework. The information model consists of a conceptual three part structure- directory, inventory, and guide- and allows implementers to interpret and apply the model to best fit their data. In addition, it facilitates management of administrative information about data and metadata by incorporating this information within the unified structure. As the diversity of data resources expands, and as users of the data and metadata increase, the model allows addition of information within the structure to accommodate new uses and views.

Data resources are described in the SEDAC Gateway and presented in three views: directory, inventory, and guide. Metadata for each serves a different function:

- Directory view - Permits the user to discover an appropriate dataset or collection of datasets;
- Inventory view - Enables the user to determine details of how to obtain a specific dataset;
- Guide view - Supplies contextual and supporting documentation that allows the user to effectively use the dataset, and may provide the ability to browse samples of the data or the data themselves.

Directory metadata describe aggregated data such as a major database, a time-series, or a collection of datasets. Included in this view are, at minimum, the title, abstract, data center, appropriate contacts, and index terms. Directory entries are indexed using controlled vocabularies for location, parameter, and discipline as well as uncontrolled general keywords. SEDAC works with members of the social science and natural science communities to develop controlled vocabularies appropriate for the description of socioeconomic data resources.

Inventory metadata are descriptions of granules, which are the smallest items managed in a data archive. Inventory metadata are not produced for all datasets described by SEDAC, but only for items for which SEDAC has permission to disseminate. Included in this view are the size, physical location (path), and, if available, the cost of the item. Inventory metadata are searchable by variables which reflect the unique content of the collection. The result of such searching is a list of "orderables" which the Gateway then makes available to the user.

Guide metadata contain detailed information designed to complement directory and inventory metadata by providing context and background information. Three kinds of guides are made available through the Gateway: dataset, organization, and thematic. Dataset guides provide an overview of the data resource and details on its source and history, access to variable lists and data dictionaries if available, and guidance in using the dataset. Organization guides give the user additional in-depth information concerning the data center which created, owns, or maintains the dataset. Thematic guides are subject oriented overviews of key topics and issues that pertain to the human interactions in global environmental change. They provide access to the full text of numerous journal articles, book chapters, conference papers, government reports, and other materials. Guides may also offer the ability to browse samples of the data or the

data themselves. Sample browse guides associated with the directory view present extracts of the data such as images, tables, or text. The inventory view may include similar browse guides about a specific subset of the data. Dynamic browse guides enable users to query, select, and view portions of the data. Guide metadata are constructed to comply with existing standards utilized in SEDAC metadata production. This augments overall SEDAC efforts to produce internally consistent metadata that is also externally standards-compliant and potentially interoperable with the requirements of other established and evolving metadata formats.

In addition to the directory, inventory, and guide metadata which are presented to the user of the Gateway, SEDAC also generates and manages metadata for internal administrative tracking of datasets and for on-site library holdings.

4.1.3 Physical Organization

Both directory and inventory metadata are stored and maintained in a relational database. From the stored directory/inventory metadata, various views or formats of the metadata can be extracted. Currently, directory entries are extracted as formatted ASCII files which are parsed and loaded onto a server for access by the Gateway. Plans for development include the ability to read directory information, in a variety of flexible views, directly from the relational database. Inventory metadata are currently read directly from the database by the Gateway.

Guide metadata are typically stored as HTML documents. The Gateway software provides the capability to access relevant guides from appropriate directory entries. Guides are also made available via SEDAC's World Wide Web Home Page.

The administrative and library metadata are currently maintained in separate, independent database structures.

SEDAC is progressing toward a vision of a unified metadatabase in which directory, inventory, guide, administrative, and library components will be integrated into one system. This will ensure complete, consistent, efficient handling of metadata across all its varieties of functions; and will also promote and facilitate interoperability between existing and evolving metadata format standards.

4.1.4 Metadata Standards

Current knowledge of standards being used in library and information management communities is essential for the development and integrity of metadata provided and distributed by SEDAC. Along with maintaining the standards already established for metadata, continued effort is required to remain knowledgeable of and involved in the wider discussions of issues and standards development for digital information resource description and dissemination. Several types of standards must be addressed including content (cataloging and indexing), interchange, and format standards. These standards continue to evolve to address the many issues associated with description and access of online resources. Complying with multiple metadata standards improves the ability of

SEDAC to provide wider access to its data and information products. The following provides a brief description of several metadata standards which constitute priority directions for SEDAC.

DIF. The Directory Interchange Format is a data structure for directory metadata developed by NASA and maintained by the Interagency Working Group on Data Management for Global Change (IWGDMGC). The format suggests metadata elements for describing data at a highly aggregated level, prescribes content values for selected elements, and provides a structure for transferring metadata among information systems. DIF-compliant directory metadata can be extracted from the relational database. The SEDAC Gateway is interoperable with other DIF-compliant systems, such as the GCMD. In addition, SEDAC participates in the IWGDMGC, which provides direction and recommendations on the development of the DIF.

GILS. The Government Information Locator Service (GILS) is an emerging standard the purpose of which is to facilitate electronic access to federal government sponsored information. SEDAC is committed to ensuring that the metadata it creates will be available to GILS users. This will be accomplished by making appropriate metadata available via a GILS compliant server and generating one or more GILS Core Records describing the holdings of that server.

MARC. MACHine Readable Cataloging (MARC) is a metadata standard used by the library community to exchange cataloging records which employ the Anglo-American Cataloging Rules for describing a myriad of resources, including machine readable data files. SEDAC is exploring interoperability with the MARC standard in order to be able to tap into a rich source of high quality, extant metadata and also to contribute metadata to the larger library community in a form usable by most library information systems. Achieving interoperability requires the review and modification of the current metadata storage schema and the development of tools to extract and import MARC records; this activity is ongoing.

Spatial Metadata Standards. Two standards efforts relating to spatial metadata are being tracked. One is the Spatial Data Transfer Standard (SDTS) which describes the content of a data transfer among GIS packages. This has been approved as a Federal Information Processing Standard by the National Institute for Standards and Technology. The other is the Content Standards for Digital Geospatial Metadata, which addresses the need to determine common terminology for spatial metadata and to define the minimum set of metadata needed to describe a spatial data resource. It was approved at the June 8, 1994 meeting of the Federal Geographic Data Committee. Executive Order 12906, "Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure," instructs federal agencies to use the standard to document new geospatial data beginning in 1995. It is SEDAC's intention to work towards being compliant with the spatial metadata content and transfer standards with respect to the digital spatial data in its collection.

4.2 National Information Infrastructure

The Internet has experienced phenomenal growth since the last version of this Information Gateway Plan. This has been fueled by the broad acceptance of the World Wide Web as a preferred mechanism for information providers and by the development of browsers such as Mosaic and Netscape, as well as search tools such as Yahoo, Lycos, Web Crawler and Harvest. In addition to these publicly popular resource discovery tools, the Government Information Locator Service (GILS) has also been established and promotes the transfer of information through the Z39.50 protocol. These advances in technology still do not recognize that vast majorities of Internet users are not within range of sufficient bandwidth and rely heavily on e-mail protocols. It is not unusual to transmit WWW home pages through e-mail. SEDAC is positioned to serve members of the user community at both ends of this spectrum. Many scientists conducting research in global change in developing countries will need to rely on rudimentary Internet connectivity, if they have any at all; others in developed regions are active users of the most modern Internet tools.

SEDAC will extend its primary search engine, Gateway, to respond to both ends of the spectrum, by enabling e-mail and web based transactions. In addition, SEDAC can increase the exposure of the resources provided by being recognized by the growing collection of resource discovery tools. These tools generally rely on information packaged in WWW home pages, hence, topical indexes, thematic guides and other web based information products will be discovered by resource seeking "knowbots" and will become cross indexed in more commonly available search tools.

SEDAC Gateway plays a unique role however. Current search tools depend upon information available in file titles and home pages. The Gateway, focuses on data relevant to Mission to Planet Earth, offers a view of data bases, and other information resources not known to other search tools. The Gateway is designed to be interoperable with the GCMD. The Information Cooperative expands the range of metadata accessible by promoting the use of the GCMD/CEOS metadata model. In addition, plans to introduce GILS interoperability will further extend the reach of the SEDAC Gateway. These steps planned to be taken by SEDAC in this contract year will further establish SEDAC's role as a Gateway to environmental resources by becoming more firmly rooted in the National Information Infrastructure.

4.3 Information Cooperative

SEDAC is participating actively in the development of the Information Cooperative, a CIESIN initiative to pull together a worldwide network of institutions that hold social science and interdisciplinary data and information. In particular, SEDAC is helping to develop the technological infrastructure that allows these institutions to make their existing electronic data catalogs and inventories more accessible to a wide range of users through a common user interface. SEDAC is also working closely with selected Information Cooperative partner organizations to help gain access to key data needed for applications development. Such institutional arrangements are vital since, in order

for SEDAC to provide operational services on a continuing basis, it must have long-term access to the updated data needed for that service.

The development of the Information Cooperative has been in progress for four years. The emphasis during this initial phase has been on well established data centers in government, non-government, and inter-governmental organizations. Table 3 lists participating organizations, relevant to SEDAC, under the heading of the eight data domains. Agreements have been developed (or will be developed) with these organizations and their metadata and data have begun to become available through the system.

During the current contract year, SEDAC will coordinate with CIESIN Information Cooperative in the following areas:

- 1) SEDAC's Role as an Information Cooperative Node;
- 2) Partnership Development Activities; and
- 3) Operational Support for Nodes of Highest Priority to SEDAC.

4.3.1 SEDAC's Role as an Information Cooperative Node

Establishing Human Dimensions Metadata Database

SEDAC will continue to strengthen its human dimensions metadata database utilizing the CIESIN Information Cooperative's strong relationships with major socioeconomic data providers. Organizations will be targeted based on the criteria such as their relevance to SEDAC's mission, their wealth of data resources, their apparent ability to develop, maintain and update metadata, and their Internet access capability. The Science Data Plan Update reprioritize the Information Cooperative partners based on the above criteria.

Distributed Data Access Initiatives

The SEDAC Information Gateway will prototype a distributed interoperable inventory metadata and data ordering capability at 3 - 4 high priority Information Cooperative partner sites during this contract year. This effort will utilize existing Gateway software developed at CIESIN. SEDAC will install the Gateway software at high priority sites with existing hardware and software facilities. SEDAC will provide a complete package of Gateway software, metadata loading and modification utilities, database schema, and a modest technical telephone support. SEDAC will modify the software and documentation so as to support the top 3 - 4 most popular UNIX platforms, to the extent necessary for the sites that will be specified in the Science Data Plan Update which is due in November 1995.

SEDAC will complete development of software, documentation, and necessary "load and modify" utilities for inventory metadata and data ordering, for the Windows NT / Oracle for IBM PC compatibles approach, which is currently being provided to nodes in

developing countries. SEDAC will attempt to establish a full inventory metadata and data ordering capability at two high priority nodes in less developed countries (to be determined), where the underlying hardware and software infrastructure have already been implemented.

SEDAC will support Interactive Gateway Query and/or Browse Access, via the Internet, for three SEDAC relevant relational databases, to be specified in the next Science Data Plan Update. Most of the work will involve reformatting data for access by the software.

Gateway Entry Nodes

Since SEDAC is committed to provide access to applications development products and other socioeconomic data to multi-national users, it is important that the data needs to be delivered at a reasonable speed of Internet access. This requires the establishment of Gateway entry nodes at regional sites as opposed to having a single point of entry. SEDAC will provide necessary technical infrastructure for establishing Gateway entry point nodes. In the contract year, Gateway entry nodes will be established in regional locations; one in Asia, one in Europe, and one in South or Central America. SEDAC users will be benefited by establishing regional Gateway Entry point nodes. For example, SEDAC users in IIASA would be greatly benefited by establishing a Gateway Entry point node in Sweden or Austria.

Supporting Partnership Development Activities

During the contract year, SEDAC will support CIESIN's Information Cooperative in supporting partnership development activities such as site coordination, potential partner identification, metadata and guide development, and setting coordinating nodes.

Site Coordination

SEDAC will continue to support site coordination and partnership development activities for access to relevant metadata and data from existing sites and will collaborate in establishing relationships with additional sites, as specified in SEDAC's Science Data Plan. This will include initial contacts, technical assessments, development of Memoranda of Understanding, and specific work plans.

Potential Partner Identification

SEDAC will review the nominations of potential Information Cooperative partners based on their relevance to SEDAC's applications development tasks. The potential partners who possess data that are required for applications development tasks and at the same time, the partners who are willing to share their data would be considered as high priority partner sites in the Science Data Plan Update. On approval of the UWG, a plan will be developed for each new site, based on its overall relevance and priority,

and a site coordinator will be assigned to work with the site for developing and maintaining the relationship.

Metadata and Guide Development

SEDAC will provide documentation, templates, existing "load and modify" utilities, professional review, and informational support and interaction for those priority nodes. SEDAC will do this from the perspective of making these nodes self supporting entities, which will do most of their own work and require minimal support over time. Furthermore, the level of support will depend on the priority each node is given in the Science Data Plan. SEDAC will not provide the same level of support to all nodes. The exact level of support will be determined by available resources and other priorities.

Coordinating Nodes

The Information Cooperative is setting up coordinating centers for specific disciplines (e.g., land use, biodiversity). These coordinating nodes identify relevant organizations and resources and coordinate access to their data by guiding them into the Information Cooperative. SEDAC will assist in identifying logical clusters of organizations and potential coordinating nodes. SEDAC will work with the Information Cooperative in identifying potential coordinating nodes and clusters relevant to SEDAC's missions. For example, SEDAC is currently working with the Council of European Social Science Data Archives (CESSDA) in order to determine an appropriate coordinating node from among CESSDA's member institutions.

4.3.2 Operational Support for SEDAC's High Priority Nodes

SEDAC's Gateway is based on the concept of linking to the diverse systems that exist and keeping pace with them as they change. This places resource demands on SEDAC's ability to maintain and expand a distributed interoperable system. Although SEDAC will only provide site coordination and technical support to specified SEDAC nodes, SEDAC will make the software and documentation for the various Gateway approaches available to the Information Cooperative, for it to provide to its wider membership.

SEDAC's operations staff will provide a modest level of telephone and e-mail consultation with regard to installation and ongoing systems maintenance to the operations staff at all SEDAC relevant distributed nodes. However, once the systems are installed, it is the responsibility of the sites to update the data and metadata.

User Requirements, Systems Development, and Testing

There is a recognition that increased inputs from users are necessary to assure adequate systems development. SEDAC customer services, data and information resource, site coordinators, and metadata staff, are key persons using the Gateway System on a daily basis. They are also in touch with external sites and users and are very aware of changing Internet tools. SEDAC will utilize them as a primary resource in developing systems requirements and interfaces and in systems testing. SEDAC will also call on key

nodes to provide inputs to ongoing iterations of requirements and testing. SEDAC and key coordinating nodes must be consciously used as prototyping test-beds for the evolving systems. End users of the system at SEDAC and at other key nodes must have early inputs into the development, software, documentation.

4.4 The SEDAC Archive

The SEDAC Local Archive acquires, archives, and disseminates data and information resources required and/or developed by the Applications Product Developments efforts. It provides internal archiving services in support of all product development work. It also archives the final products for purposes of permanent storage and for dissemination. The local archive provides archival and acquisitions support throughout the life cycle of the data. This includes data acquisition, obtaining necessary permissions for use and dissemination of the data, data ingestion, back up, quality checking, dissemination planning and implementation, integrity checking, maintenance, update, overall file management and security. The entire data management activities will be conducted in accordance with SEDAC data policies and terms of use agreements with data suppliers. The SEDAC Local Archive will acquire and ingest data that are required for various applications development activities and also social science data which could be used for value added purposes.

The local archive will disseminate SEDAC products via one or more channels as appropriate for each data and/or information product. Each major item disseminated will have a dissemination plan, which will state the users being targeted, appropriate dissemination channels and plans and strategies for notifying users that the resource is available. It will also include items such as update dates, use restrictions, and appropriate credits and disclaimers. The Local SEDAC Archive will implement access and promotion based on the dissemination plan. Notice of availability will take place largely by targeting appropriate e-mail list serves, notices and order links on the SEDAC World Wide Web Home Page, inclusion in the Gateway Catalog, and occasional exposure at appropriate conferences and public forums. Products will be disseminated so as to achieve the widest possible gateway user ship. This will frequently entail joint utilization of access through the Gateway catalog system and also via popular channels such as World Wide Web.

With the approval of the UWG and to the degree that resources permit, the local archive may also acquire, archive, and disseminate data (from the eight SEDAC science data domains addressed in Table 3), which are highly relevant to the topics and products addressed in the DADP and useful for scientists' use of these products.

The local archive will cooperate with SEDAC's Information Gateway partnership initiatives in the following manner. It will provide dissemination services, for important datasets, relevant to SEDAC's mission, when partner organizations do not have their own local means of online dissemination. It will only provide this service for datasets listed in the SEDAC Science Data Plan, as approved by the UWG. It will only provide this service for carefully selected datasets when the value-added benefit from online dissemination and/or interactive browse and query is deemed to be quite high.

Table 3. Example Information Cooperative Partners with significant holdings corresponding to SEDAC Science Data Domains.

Data Domain	Information Cooperative Partner
Land and Fresh Water Resources	• Organization of American States (OAS), Washington DC, USA
Industry and Energy	• Organization for Economic Co-Operation and Development (OECD), Paris, France
Agriculture and Food Security	• Center for Indigenous Knowledge for Agriculture & Rural Development, (CIKARD), Ames IA, USA • International Crops Research Institute for Semi-Arid Tropics (ICRISAT), Hyderabad, India • Food & Agriculture Organization (FAO), Rome, Italy • International Centre for Research in Agroforestry (ICRAF), Nairobi, Kenya
Population Dynamics	• Center for International Research (CIR), U.S. Bureau of the Census, Washington DC, USA • Social, Labor, and Institution Statistics, Statistics Canada, Ontario, Canada
Economic Activity	• Organization for Economic Co-Operation and Development (OECD), Paris, France • Statistical Office of the European Communities (Eurostat), Luxembourg
Human Attitudes, Preferences, and Behavior	• Economic and Social Research Council Data Archive (ESRC), Essex, UK • Harris Center, Institute for Research in Social Science (IRSS), Chapel Hill NC, USA • Inter-University Consortium for Political & Social Research (ICPSR), Ann Arbor MI, USA • Roper Center for Public Opinion Survey, Storrs CT, USA • Zentralarchiv für Empirische Sozialforschung (ZA), Cologne, Germany • Norwegian Social Science Data Services (NSSDS), Bergen, Norway
Policy and Institutions	• IUCN-The World Conservation Union, Gland, Switzerland • World Conservation Monitoring Centre (WCMC), Cambridge, UK • World Resources Institute (WRI), Washington DC, USA • Institute for Public Policy and Social Research (IPPSR), East Lansing MI, USA
Human and Environmental Health	• Agency for Toxic Substances & Disease Registry (ATSDR), Atlanta GA, USA • National Institute of Public Health & Environmental Protection (RIVM), Bilthoven, The Netherlands • World Conservation Monitoring Centre (WCMC), Cambridge, UK • World Health Organization (WHO), Geneva, Switzerland • Centers for Disease Control and Prevention (CDC), Atlanta GA, USA

The SEDAC local archive will also maintain a modest internal library of relevant data and bibliographic reference materials for use in Gateway and applications development initiatives.

Table 4. Information Cooperative National Partners.

Countries	Locations of Country Nodes
Estonia	Estonia National Library
Hungary	Hungary Country Node-OMIKK
Poland	Polish Academy of Sciences, University of Warsaw
Latvia	University of Latvia
China	Chinese Academy of Science, Peking University
Japan	Keio University
Thailand	Chiang Mai University

5.0 Information Gateway System Components High Level Requirements

The needs and requirements outlined above for the SEDAC Information Gateway define certain high level requirements that must be addressed by the SEDAC AOC and eventually be included in the implementation of SEDAC's information systems. These systems include four integrated components: Data and Information Access System (DIAS), Data Archiving and Distribution System (DADS), Data Processing and Analysis System (DPAS), User Services System (USS). In addition to the specific requirements for each component system, a set of general requirements also pertain to these systems:

- Develop data services that are most urgently needed by the science research programs requiring SEDAC data;
- Accommodate EOSDIS approved approaches for networking, data format, information management, and archiving systems for the interface between SEDAC and EOSDIS data centers;
- Develop a data dictionary and lexicon used for standardization of terms and definitions in query processing;
- Accommodate EOSDIS approved portability guidelines and standards for development of any new SEDAC software that interfaces with the EOSDIS community;
- Perform testing and validation of the functions and performance of DOMINO, data formats, science processing library, network management, and SEDAC specific tasks;
- Provide methods for user authentication (verification of identity) and authorization (access privileges);
- Maintain an adequate level of system security and data integrity;
 - a) open access to search of all levels of metadata;
 - b) no access to host operating system during search of metadata;
 - c) access to restricted data controlled by passwords;
 - d) media delivery of restricted data controlled by approved lists;
- Provide adequate level of documentation for all system and user functions; and
- Search, order, and distribution functions for data maintained by non-CIESIN organizations, must be defined within an Interface Control document;

5.1 Data and Information Access System

The Data and Information Access System (DIAS) allows users to locate, view, and order data and information of interest that are located within the SEDAC archive and other data and information centers who are linked to SEDAC as part of CIESIN's Information Cooperative with systems compatible with DIAS. Additional resources include (but are not limited to) data and information available through the Global Change Master Directory and NASA's EOSDIS system, and other resource servers publicly available on the Internet. DIAS includes two subsystem components, CIESIN Gateway and Information Guides.

Gateway is the delivery mechanism for directory and inventory level metadata as well as the interface to the SEDAC data ordering system. Gateway is a critical resource to facilitate searching and identification of relevant global change physical and socioeconomic datasets.

Information Guides include additional metadata information organized around datasets, organizations, or themes such as land use, remote sensing, and ozone depletion. Guides provide important contextual and background information to facilitate use and understanding of global change data available through SEDAC. Because of the textual nature of the guides, a hypertext interface is a useful access mechanism for this service. Other requirements include:

In order to meet the user needs described above, DIAS must provide these functionalities:

- Provide mechanisms to accept and process requests via telephone, postal service, or electronic mail;
- Develop capability to achieve at least Level 2 catalog interoperability between the SEDAC and cooperating archive centers;
- Work toward Level 3 catalog interoperability between the SEDAC and cooperating archive centers;
- Populate and make available a directory that describes the data holdings of the SEDAC and cooperating archive centers;
- The directory must be functionally compatible with the GCMD (i.e. SEDAC systems can search directories that contain standard GCMD entries.)
- Be able to transfer directory interchange format entries (DIF's) with the Global Change Master Directory (GCMD);
- Provide access catalogs and guides that provide textual information, including detailed descriptions of data across SEDAC and cooperating archive centers;
- Maintain online metadata for SEDAC and cooperating archive centers;

- Provide online search of metadata for SEDAC and cooperating archive centers;
- Provide online and offline distribution of metadata for SEDAC and cooperating archive centers;
- Accommodate EOSDIS IMS recommended standards for metadata search for any information searches that originate from EOSDIS user interfaces;
- Provide standard procedures for submitting orders for data products held at SEDAC and cooperating archive centers;
- Utilize a common user interface that supports both alphanumeric and GUI style user interfaces including a WWW interface;
- Include an e-mail interface for low bandwidth connectivity
- Provide a data search capability to locate datasets of interest residing at SEDAC and cooperating archive centers by common attributes. (e.g., spatial location, coverage, and time.)
- Provide a data search capability to locate data granules of interest within specific datasets by data granule attributes (e.g., sea surface temperatures.)

Support queries to and from Z39.50 based systems;

- Provide a browse capability for improving users' data selection;
- Collect statistics for long term monitoring of the distributed system performance;
- Support the evaluation of spatial search techniques and the development of spatial search standards;
- Provide online and offline data order and distribution of data order and distribution of data use documentation;
- Access to directory and inventory metadata;
- Access to Data Archive and Distribution System (DADS); and
- Support for heterogeneous systems.

5.2 Data Archive and Distribution System

The Data Archive and Distribution System (DADS) provides archive management for all SEDAC holdings, process orders for user selected data and software, deliver data and software to users, and ingest new data and software into the archive. DADS provides the infrastructure that supports the SEDAC archive, and may also be utilized at other sites with compatible systems and infrastructure. The requirements for DADS include:

- Ingest algorithms, ancillary data, and metadata from SEDAC PGS into SEDAC DADS;
- Accommodate EOSDIS standard procedures for submission of any data, correlative data, and metadata to EOSDIS archive systems;
- Receive data on SEDAC approved magnetic tape or CD-ROM, or via electronic transmission;
- Provide mechanisms to convert data inputs from paper to electronic form prior to ingest into SEDAC archives;
- Provide data submission procedures for individuals or organizations submitting data products for SEDAC archival storage and distribution;
- Distribute data in response to DIAS requests;
- Make pre-generated browse products available to support search and order;
- Distribute previously archived SEDAC data, correlative data, algorithms, and documentation to SEDAC users within a month of the order request;
- Distribute data from cooperating archive centers according to procedures outlined in formal agreements between SEDAC and those centers;
- Accommodate the distribution of data in approved SEDAC formats and media;
- Accommodate the distribution of data via electronic mail;
- Provide mechanisms to print and distribute hard-copy versions of "readable" archive contents;
- Bring SEDAC data to minimum management standards (e.g., resident on good quality media with inventory metadata);
- Provide different levels of SEDAC archive access performance through a tiered system of disk, CD-ROM, and tape;
- Exchange with EOSDIS and related centers, data products, algorithms, ancillary data, and metadata for processing or archiving;
- Support the collection, assembly, and publication of datasets on CD-ROM for distribution to users;
- Maintain accurate and complete online and offline documentation for users of data;
- Insure recoverability of SEDAC archive data according to SEDAC operational plan;

- Provide users with information about recoverability of data at SEDAC and SEDAC-related archives;
- Provide mechanisms for audit and traceability of all data transfer;
- Accommodate the use of HDF data format and packaging for:
 - a) Users submitting data to SEDAC archives
 - b) Distribution of data to facilities requiring HDF formatted data
 - c) Earth science data stored in SEDAC archives;

5.3 Data Processing and Analysis System

The Data Processing and Analysis System (DPAS) provides computational tools and support for product generation, scientific analysis, application and tool development, and human science-physical science integration studies. User needs dictate that DPAS must meet four primary functions:

- Acquire and develop software tools useful for:
 - a. Geo-referencing of socioeconomic data
 - b. Development of base maps for GIS
 - c. Mapping of socioeconomic data
 - d. Data conversion and conditioning
 - e. Interactive data browsing and extraction
 - f. Other visualization and analysis capabilities;
- Develop data products using these tools;
- Provide tools for users to use on their local equipment; and
- Provide interactive integration and analysis services using tools on SEDAC equipment.

5.4 User Services System

The User Services System provides online and traditional assistance in data access, ordering, and analysis. The primary component of USS is a staff of social and physical science experts. Supporting technical systems consist of:

- Provide a help desk;
- Provide a bulletin board;
- Assist users in using SEDAC systems;
- Assist users to identify and locate datasets;

- Help users understand "phased" levels of service (e.g., when is a particular archive or catalog service operating at full capacity as opposed to initial, alpha, or beta levels.)
- Coordinate with other centers to assist users;
- Compile user service statistics;
- Provide system usage reports;
- Provide system demonstrations;
- Provide 40 hour prime time staffing (U.S. Eastern time zone);
- Provide limited support services for "off" hours. (i.e. after-hour requests and requests during business hours in other time zones);
- Track revisions of data;
- Track validation of data;
- Develop and maintain a dictionary of unique services; and
- Provide internal and external training.

References

- CEES (Committee on Earth and Environmental Sciences). 1992. *Reducing the Impacts of Natural Hazards: A Strategy for the Nation*. CEES Subcommittee on Natural Disaster Reduction.
- CEES (Committee on Earth and Environmental Sciences). 1993. *Our Changing Planet: The FY 1993 U.S. Global Change Research Program*.
- CEES (Committee on Earth and Environmental Sciences). 1994. *Our Changing Planet: The FY 1994 U.S. Global Change Research Program*.
- CENRR (Committee on Environment and Natural Resources Research). 1995. *Our Changing Planet: The FY 1995 U.S. Global Change Research Program*.
- CIESIN (Consortium for International Earth Science Information Network). 1992. *Pathways of Understanding: The Interactions of Humanity and Global Environmental Change*. Saginaw MI: CIESIN.
- Croze, H. and V. Vandeweerd. 1991. Monitoring and data management technology for environmental assessment. In: *Proceedings of the Environmental Information Forum* (Montreal, Canada, May 1991).
- Jacobson, H.K. and M.F. Price. 1990. *A Framework for Research on the Human Dimensions of Global Environmental Change*. HDP Report 1.
- LUCC (Land-Use and Land-Cover Change) Core Project Planning Committee. 1995. *Science Plan for Land-Use and -Cover Change*. In press.
- NASA (National Aeronautics and Space Administration). 1995a. *Mission to Planet Earth Strategic Enterprise Plan 1995-2000*. Washington DC: NASA.
- NASA (National Aeronautics and Space Administration). 1995b. *Proceedings, EOSDIS Potential User Model Development Effort, Preliminary Draft*. Washington DC: NASA Office of Mission to Planet Earth, Operations, Data and Information Division.
- NRC (National Research Council). 1990. *Research Strategies for the U.S. Global Change Research Program*. Washington DC: National Academy Press.
- NRC (National Research Council). 1991. *Solving the Global Change Puzzle: A U.S. Strategy for Managing Data and Information*. Washington DC: National Academy Press.
- NRC (National Research Council). 1992. *Global Environmental Change: Understanding the Human Dimensions*. Washington DC: National Academy Press.
- NRC (National Research Council). 1994a. *Science Priorities for the Human Dimensions of Global Change*. Washington DC: National Academy Press.
- NRC (National Research Council). 1994b. *Memoranda from a Workshop on Research Needs and Modes of Support for the Human Dimensions of Global Change*. Washington DC: Committee on the Human Dimensions of Global Change.
- NRC (National Research Council). 1995. *A Review of the U.S. Global Change Research Program and NASA's Mission to Planet Earth/Earth Observing System*. Washington DC: National Academy Press.
- NSF (National Science Foundation). 1988. *Mosaic* 18(3/4):1.

Yohe, G., and K. Segerson. 1992. *Economic Data and the Human Dimensions of Global Environmental Change: Creating a Data Support Process for an Evolving Long Term Research Program*. HDP Report 4.

Acronyms and Abbreviations

AGCI	Aspen Global Change Institute
AOC	Architecture and Operational Concept
ASCII	American Standard Code for Information Interchange
ATSDR	Agency for Toxic Substances & Disease Registry
CD-ROM	Compact Disk - Read-Only Memory
CDC	Centers for Disease Control and Prevention
CEES	Committee on Earth and Environmental Sciences
CENRR	Committee on Environment and Natural Resources Research
CESSDA	Council of European Social Science Data Archives
CEOS	Committee on Earth Observations Satellites
CFC	chlorofluorocarbon
CGIAR	Consultative Group for International Agricultural Research
CIESIN	Consortium for International Earth Science Information Network
CIKARD	Center for Indigenous Knowledge for Agriculture & Rural Development
CIR	Center for International Research
CIRAN	Centre for International Research in Advisory Networks
COTR	Contracting Officer Technical Representative
DAAC	Distributed Active Archive Center
DADP	Data and Applications Development Plan
DADS	Data Archiving and Distribution System
DIAS	Data and Information Access System
DIF	Directory Interchange Format
DPAS	Data Processing and Analysis System
EDC	EROS Data Center
EOS	Earth Observing System
EOSDIS	Earth Observing System Data and Information System
ESRC	Economic and Social Research Council Data Archive
EUROSTAT	Statistical Office of the European Communities
FAO	Food and Agriculture Organization of the United Nations
FTP	File Transfer Protocol
GCDIS	Global Change Data and Information System
GCMD	Global Change Master Directory
GILS	Government Information Locator Service
GIS	Geographic Information System
GRID	Global Resource Information Database (UNEP)
GSFC	Goddard Space Flight Center
GUI	Graphical User Interface
HDF	Hierarchical Data Format
HDP	Human Dimensions of Global Environmental Change Programme
HDP/DIS	Human Dimensions Programme Data and Information System

HTML	HyperText Markup Language
IC	Information Cooperative
ICPSR	Inter-University Consortium for Political & Social Research
ICRAF	International Centre for Research in AgroForestry
ICSU	International Council of Scientific Unions
IGBP	International Geosphere-Biosphere Programme
IGP	Information Gateway Plan
IMS	Information Management System (EOSDIS)
IPPSR	Institute for Public Policy and Social Research
IRC	Information Resource Center
IRSS	Institute for Research in Social Science
ISSC	International Social Science Council
IUCN	International Union for Conservation of Nature
IWGDMGC	Interagency Working Group on Data Management for Global Change
MARC	MAchine-Readable Cataloging
MS	Microsoft Corporation
MTPE	Mission to Planet Earth
NASA	National Aeronautics and Space Administration
NII	National Information Infrastructure
NRC	National Research Council
NSSDS	Norwegian Social Science Data Services
OAS	Organization of American States
OCLC	Online Computer Library Center
OECD	Organization for Economic Co-Operation and Development
PGS	Product Generation System
RIVM	National Institute of Public Health & Environmental Protection
SDP	Science Data Plan
SDTS	Spatial Data Transfer Standard
SEDAC	Socioeconomic Data and Applications Center
UK	United Kingdom
UN	United Nations
UNCED	United Nations Conference on Environment and Development
UNEP	United Nations Environment Programme
USGCRP	U.S. Global Change Research Program
USS	User Services System
UV	ultraviolet radiation
UWG	User Working Group (SEDAC)
V(0)	Version 0
WCMC	World Conservation Monitoring Centre
WHO	World Health Organization
WRI	World Resources Institute
WWW	World Wide Web
ZA	Zentralarchiv für Empirische Sozialforschung

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE October 1995	3. REPORT TYPE AND DATES COVERED Annual - Oct 1, 1994-Sept 30, 1995		
4. TITLE AND SUBTITLE SEDAC Information Gateway Plan V(1)			5. FUNDING NUMBERS	
6. AUTHOR(S) Robert S. Chen Principal Investigator - compiler				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Consortium for International Earth Science Information Network 2250 Pierce Rd. University Center, MI 48710			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration, Goddard Space Flight Center Greenbelt Road, Building 12 Greenbelt, MD 20771			10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION/AVAILABILITY STATEMENT			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) This report provides a summary of the activities performed in connection with NASA/GSFC Letter Contract NAS5-3263 for the period shown.				
14. SUBJECT TERMS			15. NUMBER OF PAGES 58	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT	



2250 Pierce Road University Center, MI 48710 USA
Phone: (517)797-2700 Fax: (517)797-2622

1968 Green Road P.O. Box 134003 Ann Arbor, MI 48113 USA
Phone: (313)663-5650 Fax: (313)663-6622

1825 K Street, N.W., Suite 805 Washington, D.C. 20006 USA
Phone: (202)775-6600 Fax: (202)775-6622